



Emiliano Bobadilla Franco 198554

Ivan Salvador Razo Zapata

Visualización de Información

Curso Otoño 2022

Introducción

En la actualidad México sufre de un elevado número de carencias, pero ninguna de ellas se compara con la profunda crisis de pobreza que se ha experimentado en la época contemporánea. La medición oficial de la pobreza en el país la realiza el Consejo Nacional de Evaluación de la Política de Desarrollo Social (Coneval), y según su estudio más reciente el 43.9% de la población mexicana vive en pobreza, y el 8.5% de la población, es decir, casi 6.6 millones de personas viven en pobreza extrema. (Coneval, 2022). El problema de la pobreza va más allá de lo económico, y es por eso que, México, es el primer país en el mundo en tener una medición que utiliza un sistema multidimensional, así, la pobreza que padecen las personas no solo se refleja en el ingreso, sino también en la carencia de derechos sociales que debe otorgar el Estado mexicano a sus ciudadanos. Por una parte, la dimensión del bienestar económico comprende la capacidad adquisitiva de bienes y servicios asociados a las necesidades de la población mediante el ingreso, y, por otra parte, la dimensión social, incluye 6 indicadores de acceso a: servicios de salud, seguridad social, servicios básicos, calidad y espacios de la vivienda, educación y alimentación.

El trabajo expuesto en el presente se realizó con la intención de lograr una mejor comprensión de la problemática, con el principal objetivo de demostrar la multidimensionalidad de la pobreza, y el uso de la ubicación geográfica como herramienta para la medición y predicción de la misma. En las siguientes páginas se expone la descripción del trabajo realizado, el análisis visual de la problemática que incluye las visualizaciones obtenidas y los hallazgos más relevantes de los datasets examinados, y, finalmente, las conclusiones sobre la investigación que se llevó a cabo, así como una propuesta de proyecto para comenzar a combatir la agravante situación de pobreza del país a través de la ciencia de datos.

Descripción del trabajo

El proyecto se realizó en dos fases: investigación/recaudación de datos e implementación de visualizaciones.

Investigación/ recaudación de datos

Los datasets utilizados fueron tomados de la Encuesta Nacional de Ingresos y Gastos de los Hogares (ENIGH) realizada por el Instituto Nacional de Estadística y Geografía (INEGI), en su más reciente edición llevada a cabo del 21 de agosto al 28 de noviembre de 2020. El objetivo de la ENIGH es: *“proporcionar un panorama estadístico del comportamiento de los ingresos y gastos de los hogares en cuanto a su monto, procedencia y distribución; adicionalmente, ofrece información sobre las características ocupacionales, sociodemográficas y acceso a alimentación de los integrantes del hogar, así como las características de la infraestructura de la vivienda y el equipamiento del hogar”*.

Se tomaron tres tabulados de la encuesta para construir el dataset principal con el que se implementó el proyecto:

- Tabulados básicos
- Tabulados por entidad federativa
- Tabulados de hogares y viviendas

El dataset principal cuenta con información de las 32 entidades federativas que componen México con el fin de realizar el análisis deseado a partir de la distribución geográfica del país. De cada entidad federativa se tiene:

- Población
- Pib per capita
- Población indígena

- Coordenadas del estado
- Número de viviendas
- Datos de los 6 indicadores de la dimensión social
 - Falta de alimentos
 - Rezago educativo
 - Acceso a servicios de salud
 - Acceso a servicios básicos (se tomó como medidor el agua potable)
 - Afiliación a instituciones de seguridad social
 - Carencia de vivienda (se tomó como medidor materiales de construcción)

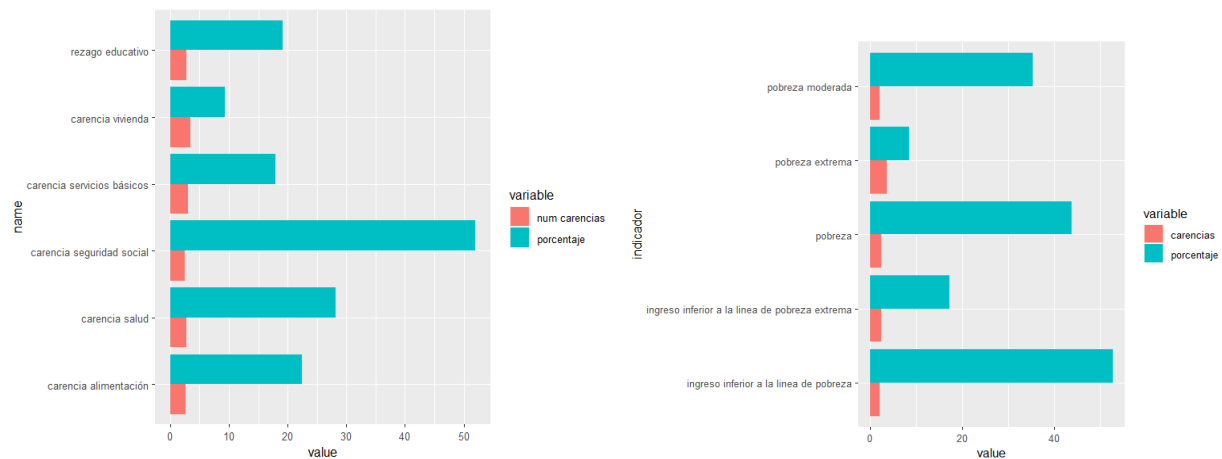
Implementación de visualizaciones

Una vez obtenidos y ordenados los datos se llevó a cabo la implementación de visualizaciones para obtener un análisis visual de la problemática de la pobreza. Todo el código se realizó en Rstudio. Se utilizaron las librerías:

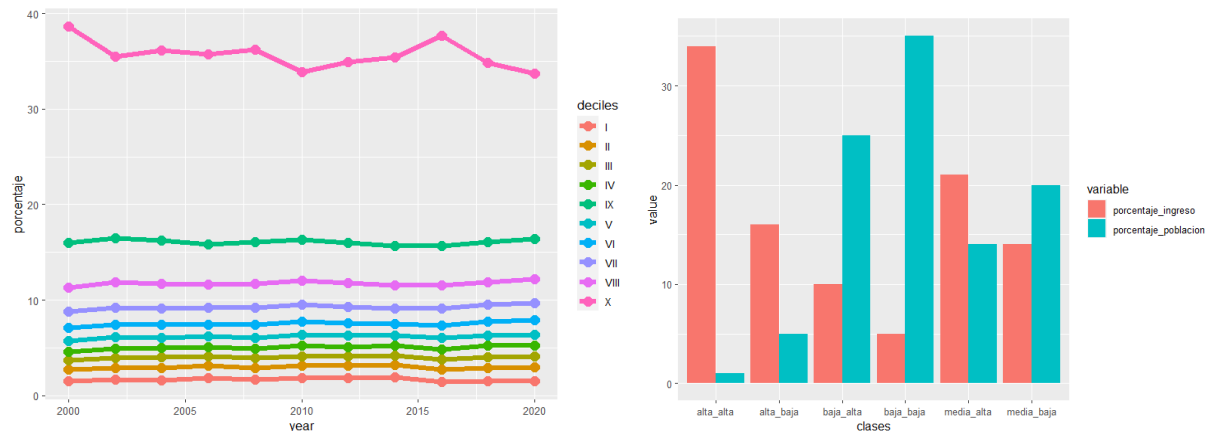
- Rjson
- Leaflet
- Leaflet.extras
- Multinet
- Ggplot2
- Dplyr
- Gridextra
- Shiny
- Shinydashboard

Se implementaron distintos tipos de gráficos: bar charts, leaflets, choropleths y redes multicasas. Finalmente para la presentación de las visualizaciones obtenidas se hizo uso de la herramienta Shiny. Con Shiny se implementó el dashboard que contiene la presentación final del trabajo realizado, y cuenta con interactividad para manipular algunas visualizaciones.

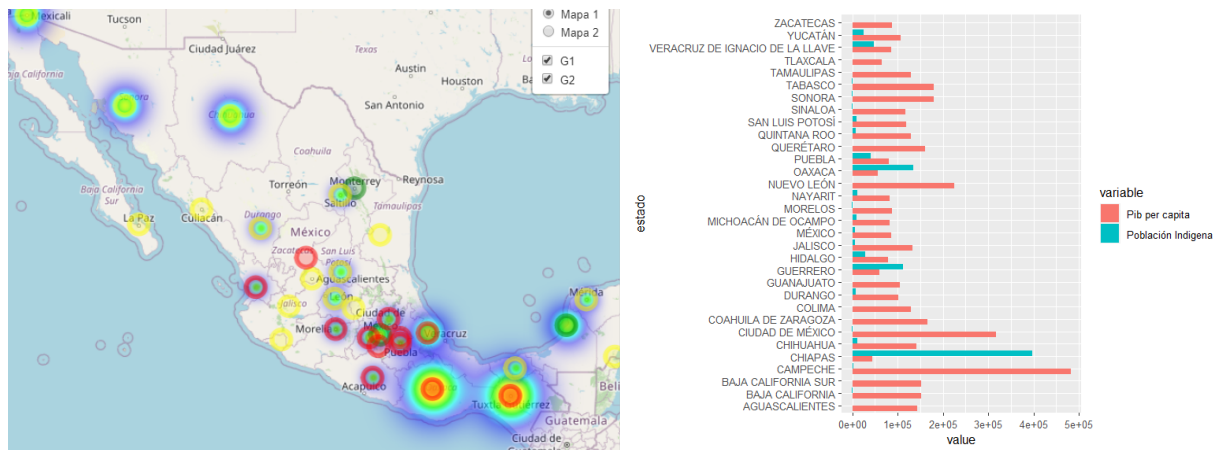
Análisis visual



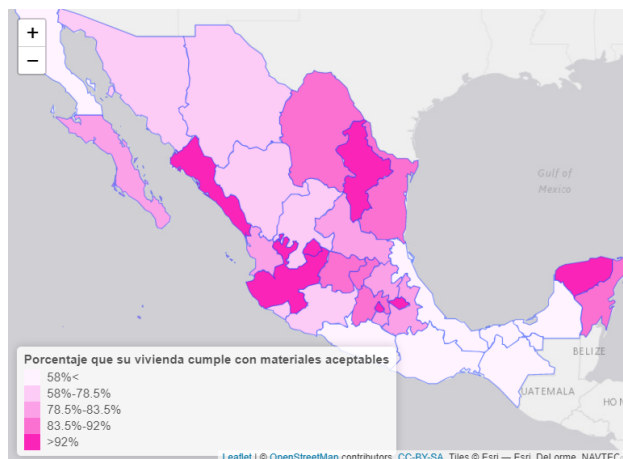
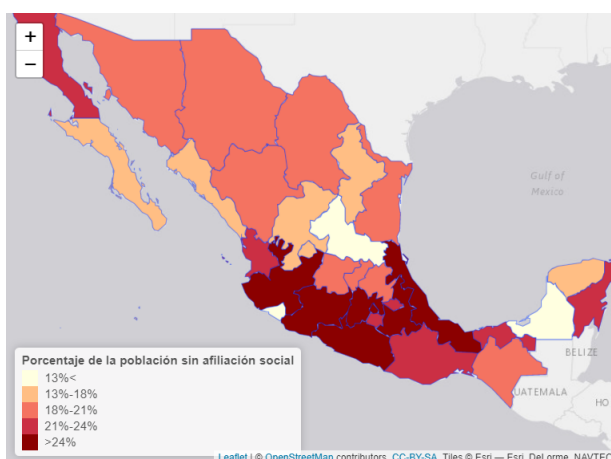
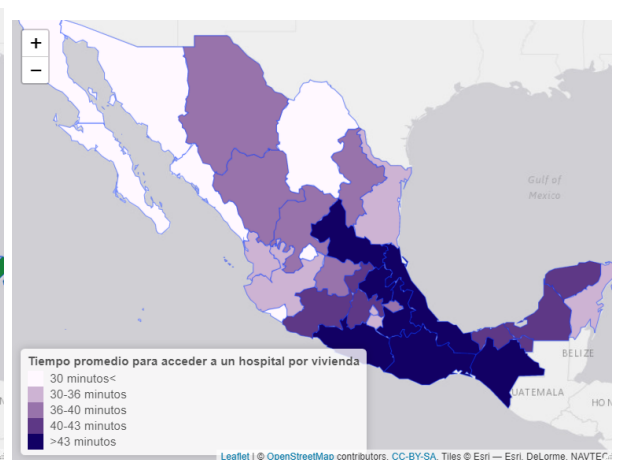
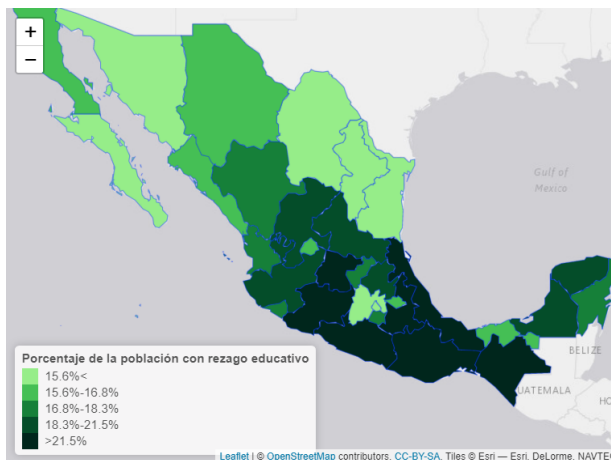
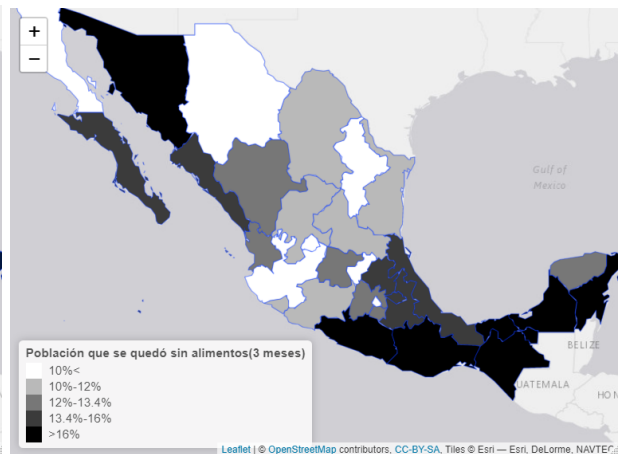
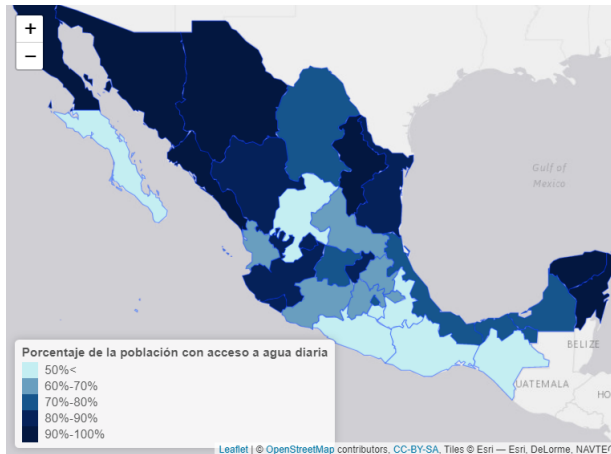
Se realizaron dos bar charts sobre los indicadores de la pobreza para contextualizar el aspecto multidimensional de la misma. A la izquierda se encuentra el porcentaje de la población con carencias de derechos sociales y a la derecha indicadores económicos. En ambos se incluye el número promedio de carencias en rojo de cada grupo que se muestra en las gráficas.



Explorando el aspecto económico, a la izquierda se muestra el comportamiento de la distribución del ingreso total generado en México a través de los años. Esta distribución se cataloga de acuerdo con los deciles de la CONEVAL. A la derecha se complementa el análisis de la distribución de la riqueza con un bar chart que toma las 6 principales clases (calculado a partir de los deciles) y compara el porcentaje de la población que pertenece a dicha clase social y el porcentaje de ingreso que este mismo porcentaje recibe, obteniendo un contraste impactante.

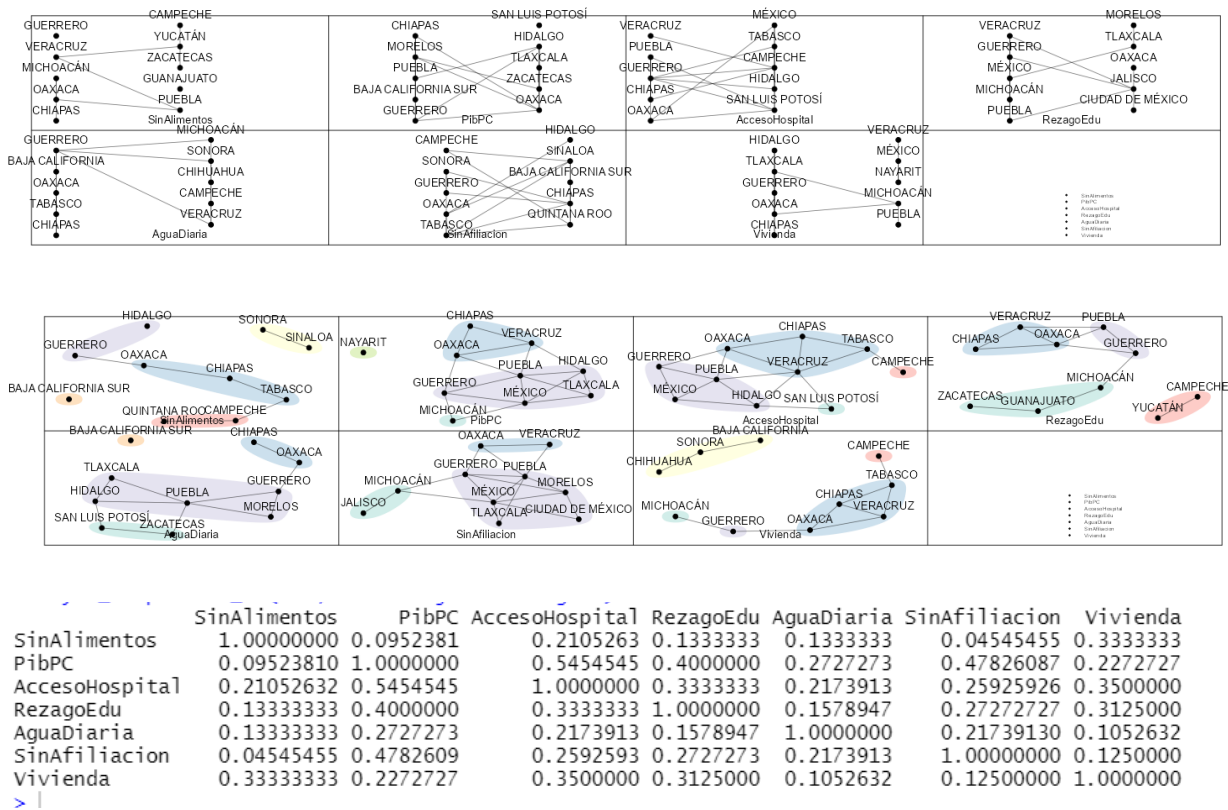


Para realizar el análisis geográfico de la variable económico de la medición multidimensional de la pobreza, se tomó como ejemplo la relación que existe entre el pib per cápita promedio que se tiene entre la población de cada entidad federativa y el número de población indígena de las mismas. Los resultados obtenidos se pueden apreciar en el leaflet de la izquierda y el bar chart de la derecha: la distribución de la población indígena en el país tiene una relación directa con los ingresos de los estados. A mayor concentración de población indígena, menor es el pib per cápita promedio, y viceversa, a menor concentración de población indígena, mayor es el pib per cápita. Esto nos habla de una clara situación de la configuración geográfica del país, mostrando una relación más que evidente.



Pasando al análisis geográfico de la medición que tiene que ver con el acceso de la población a derechos sociales otorgados por el estado, se realizaron 6 choropleths con relación a cada uno de estos indicadores: educación, alimentación, vivienda, servicios de salud, seguridad

social y servicios básicos dentro de la vivienda. Los resultados visuales son bastante claros en los 6 mapas presentados: las entidades federativas del sur y algunas del norte son una constante en cada una de las carencias por acceso a derechos sociales. Se muestra cómo ciertas zonas específicas del país son las más afectadas por la pobreza, lo que refuerza la idea de que la problemática analizada tiene una relación muy profunda con la configuración del país, el cual tiende a tener tendencias centralizadoras.



Finalmente, se realizó un análisis de redes multicapa donde hay 7 capas que representan las variables analizadas en el trabajo: por la parte económica el pib per cápita, y por la parte social los 6 indicadores sociales previamente mencionados. Los vértices representan las 10 entidades federativas que peor se desempeñan en cada uno de los 7 rubros. En la segunda gráfica se muestra una detección de comunidades muy interesante que apoya la idea mencionada anteriormente sobre la relación geográfica y la configuración de México, pues cada arista

representa estados que colindan uno con el otro. Finalmente, se muestra una matriz que representa la comparación de las 7 capas, obteniendo como resultado valores numéricos que van del 0 al 1. El valor 0 significa que no hay ninguna relación entre la comparación capa a capa, mientras que 1 significa que las capas son idénticas. Esto lo que nos muestra es cómo se repiten las entidades federativas en cada capa, reforzando la conclusión presentada anteriormente a partir de un simple vistazo a los mapas, pero ahora con números.

Conclusiones

Al analizar los resultados preliminares de la investigación expuesta en el presente a través de los datos y las visualizaciones, se puede confirmar la hipótesis propuesta al inicio del planteamiento del proyecto: existen poblaciones específicas que son las más afectadas por la configuración del país, en casi todos los aspectos, tal es el caso de Chiapas, Oaxaca y Guerrero, que son el común denominador de los estados con mayores carencias. Esto solo reafirma la necesidad de un proyecto para lograr atacar estas carencias específicas en zonas particulares del país.

Muchos de los datos expuestos en el trabajo presente son solo la punta del iceberg, y brindan información muy superficial sobre la situación específica de las personas que padecen de pobreza. Sin embargo, siendo completamente consciente de la urgencia para revertir la situación a corto plazo se puede implementar un proyecto que se base en la conexión entre los programas sociales del gobierno y los focos de atención más graves que existen en el país. Al mismo tiempo se pueden lograr predicciones sobre poblaciones o zonas puntuales con las que no se cuenta información a partir de los datos ya existentes de familias con situaciones similares en zonas cercanas, ya que como se presentó anteriormente la localización geográfica es un gran predictor para la pobreza.

Al existir esta deficiencia de datos específicos para lograr una mejor distribución de los programas sociales en el país, a mediano plazo se buscaría la recaudación precisamente de esta información para construir una red vasta de datos que permita atacar inmediatamente los focos que se van presentando en el país.

Finalmente, a largo plazo se busca que el proyecto propuesto a partir de la ciencia de datos logre tener un impacto en la política pública, tanto a nivel local como federal, proponiendo a partir de la recaudación de los datos la creación de cada vez más programas sociales que atiendan las necesidades de la población y eliminando aquellos que demuestran ser ineficientes.