

Análisis de Factores Demográficos en la Tasa de Mortalidad y Natalidad de la República Mexicana.

Diego Zúñiga

Febrero 2024

1 Pregunta de Investigación

¿Se pueden identificar los factores demográficos y sociales que afectan a las tasas de mortalidad y natalidad en las distintas entidades de México desde 2020?

2 Antecedentes

En el ámbito demográfico, tanto la mortalidad como la natalidad son fuerzas fundamentales que influyen en la dinámica poblacional. Los nacimientos y las defunciones, como pilares de esta dinámica, delinean la evolución de una población. La reducción de la mortalidad, resultado de avances médicos y mejoras en las condiciones de vida e higiene, ha sido un impulsor clave de la explosión demográfica, contribuyendo significativamente a la transición demográfica, especialmente en países en desarrollo.

La mortalidad exhibe variaciones considerables entre países, regiones, estratos sociales y grupos culturales, lo que subraya la importancia de examinar tanto las tendencias generales como las particularidades de grupos específicos y las causas principales de muerte. Esta disparidad se refleja también en diferencias de género y edad: los hombres suelen experimentar tasas de mortalidad más altas que las mujeres en la mayoría de las edades, y el riesgo de mortalidad es particularmente alto durante la infancia, disminuyendo antes de aumentar nuevamente en la vejez.

La mortalidad se clasifica en dos tipos principales: endógena y exógena. La primera está relacionada con causas intrínsecas del individuo, como malformaciones congénitas y enfermedades relacionadas con el envejecimiento, mientras que la mortalidad exógena es el resultado de factores ambientales y sociales, prevenible y tratable mediante avances médicos y acceso a servicios básicos como salud, educación y alimentación.

Por otro lado, la tasa de natalidad, como indicador crucial para el crecimiento poblacional, impacta diversos aspectos de la sociedad, desde la economía hasta la salud y el medio ambiente. Esta tasa, definida como el número de nacimientos vivos por cada 1,000 personas en una población durante un año determinado, está influenciada por factores sociales, culturales y económicos.

Los niveles de ingresos y educación juegan un papel crucial en las tasas de natalidad. En países desarrollados, un aumento en estos niveles se asocia con una disminución en las tasas de natalidad, ya que las mujeres con mayores niveles educativos y económicos tienen más recursos para planificar sus familias. Además, el acceso a servicios de salud de calidad, incluida la planificación familiar, también influye en las tasas de natalidad, ya que las mujeres que tienen acceso a estos servicios tienden a tener menos hijos.

Las creencias y prácticas culturales también afectan la decisión de tener hijos, ya que algunas culturas valoran familias numerosas, mientras que otras promueven familias más pequeñas. Las políticas gubernamentales, como licencias de maternidad, subsidios de cuidado infantil e incentivos fiscales, también pueden influir en las tasas de natalidad, alentando o desalentando a las parejas a tener hijos.

En cuanto al acceso a la atención médica, este juega un papel crucial en las tasas de mortalidad, y su análisis revela una serie de factores interrelacionados que influyen en esta dinámica. Las disparidades geográficas, socioeconómicas, la calidad de la infraestructura médica, la telemedicina, las barreras culturales y lingüísticas, la cobertura de seguro médico y el acceso oportuno a especialistas son aspectos clave que determinan la disponibilidad y calidad de la atención médica, y por ende, las tasas de mortalidad.

3 Justificación

La justificación para explorar factores no biológicos en relación con las tasas de mortalidad y natalidad, además de considerar su impacto económico y la necesidad de una planificación y gestión eficiente de recursos, radica en la búsqueda de un enfoque innovador y holístico para abordar estos temas demográficos.

En cuanto al impacto económico, comprender cómo factores no biológicos afectan las tasas de mortalidad y natalidad es crucial para entender sus implicaciones en la economía. Estos factores pueden influir en la fuerza laboral, la productividad y los costos asociados con la salud y la educación, lo que subraya la importancia de investigar más allá de los aspectos biológicos para desarrollar estrategias económicas y sociales efectivas.

En lo que respecta a la planificación y gestión de recursos, explorar factores no biológicos ofrece una perspectiva más amplia sobre las necesidades y desafíos demográficos. Esto permite una planificación más integral que no solo se centre en la atención médica y la biología, sino también en aspectos como la educación, la cultura, el acceso a servicios sociales y las políticas gubernamentales.

Además, buscar un enfoque que vaya más allá de los factores biológicos refleja un deseo de innovación y originalidad en la investigación demográfica. Al explorar aspectos como la cultura, la economía, las políticas y otros factores no biológicos, se puede obtener una comprensión más completa y contextualizada de las tasas de mortalidad y natalidad, lo que puede conducir a soluciones más creativas y efectivas para abordar este tema complejo.

4 Objetivos

- Determinar qué variables demográficas tienen una influencia significativa en las tasas de mortalidad y natalidad, proporcionando una comprensión más profunda de los impulsores de estos indicadores en México.
- Analizar patrones y tendencias en las tasas de mortalidad y natalidad a través de diferentes segmentos de la población y regiones geográficas, destacando las variaciones y disparidades existentes.
- Sintetizar los hallazgos de la investigación en recomendaciones prácticas y basadas en evidencia para responsables de políticas, profesionales de la salud y otros actores interesados en mejorar las tasas de mortalidad y natalidad en México.

5 Métodos

5.1 Bases de Datos

Con la finalidad de alcanzar los objetivos propuestos, se recopilaron 17 variables en total proporcionadas por el INEGI, con respecto a cada estado de la República. Estas variables son referentes al año 2020 y podemos encontrarlas en su sitio web oficial: <https://www.inegi.org.mx/temas/>.

5.2 Análisis Descriptivo

Se cuenta con 13 variables predictoras y 2 de respuesta (natalidad y mortalidad); para las predictoras la mayoría son demográficas y por estado.

1. Población total por estado. (poblacionT)
2. Población alfabeta por estado. (PobAlfabeta)
3. Defunciones por estado. (defunciones)
4. Población con discapacidad, con limitación en la actividad cotidiana y con algún problema o condición mental, por estado. (discapacidad)
5. Divorcios registrados, tanto administrativo como judicial, por estado. (divorcios)
6. Inmigrantes por estado. (inmigrantes)
7. Emigrantes por estado. (emigrantes)
8. Población con licenciatura por estado. (licenciatura)
9. Matrimonios registrados por estado. (matrimonios)
10. Nacimientos por estado. (nacimientos)
11. Víctimas de accidentes de tránsito terrestre por estado. (victimasAcc)
12. Población en situación de pobreza, tanto moderada como extrema, por estado. (pobreza)
13. Población que profesa algún tipo de religión. (religion)
14. Población afiliada a algún servicio de salud. (afiliadosSalud)
15. Tasa de casos nuevos de enfermedades sobre trastornos mentales y del comportamiento seleccionados por estado. (transtornos)

Decidimos hacer una gráfica de pares para observar tendencias y patrones. Con base en el orden de variables presentado, tenemos el gráfico de pares en la figura 1. Estandarizamos las variables para obtener una mejor interpretación de las correlaciones. Lo que se puede notar es que la variable Inmigrantes, Emigrantes y Población con Transtornos parecen no tener correlación con las variables restantes. De ahí en fuera se ve cierta correlación entre cada par de variables. (línea 195 del script)

Además podemos observar que en la mayoría de las variables existen posibles outliers. Para el ajuste de nuestro modelo no utilizaremos datos estandarizados.

Obteniendo su vector de medias (Figura 2) observamos que "Religión" tiene la mayor magnitud promedio entre todas las variables consideradas, mientras que "Trastornos" tiene la menor

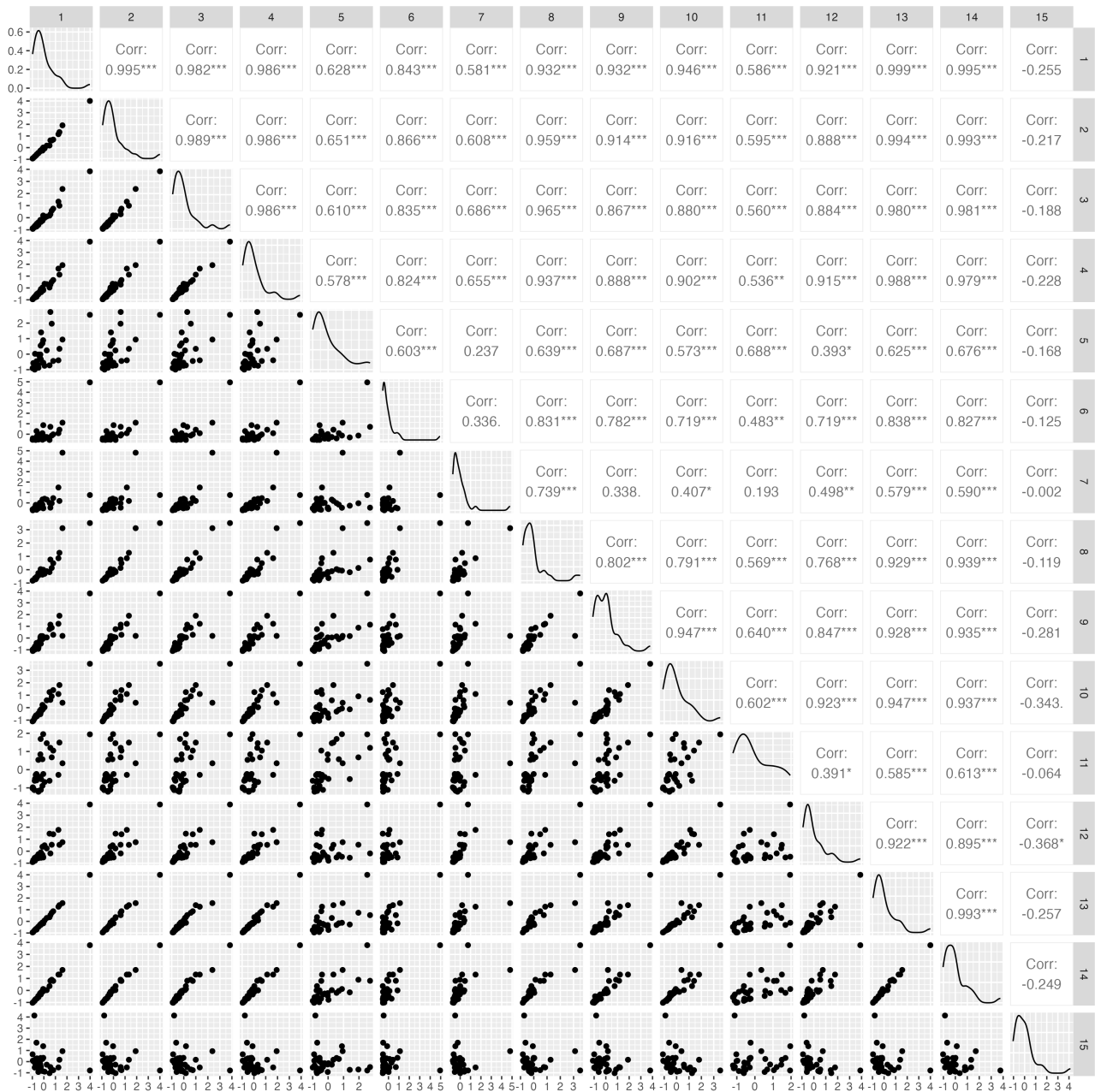


Figure 1: Gráfico de Pares de las 15 variables estandarizadas

magnitud promedio. Respecto a las desviaciones estándar, observamos que "PoblacionT" muestra una desviación estándar de aproximadamente 3,278,009, indicando una amplia dispersión respecto a su media; observación que se esperaba. En contraste, "Transtornos" exhibe una desviación estándar mucho más baja, alrededor de 121.08, sugiriendo una menor variabilidad en comparación con otras variables como "PobAlfabeta".

6 Modelo Estadístico

6.1 Explicación

En el marco de este estudio, se espera ajustar un modelo lineal generalizado, con el fin de percibir las relaciones entre las variables predictoras y las 2 variables de respuesta.

	Media	Desviación Estándar
poblacionT	3937938.2500	3.278009e+06
PobAlfabeta	2791272.5938	2.414206e+06
defunciones	33537.6562	3.102828e+04
discapacidad	1302381.7500	1.093580e+06
divorcios	2898.0938	2.376110e+03
inmigrantes	675373.8438	9.127455e+05
emigrantes	675373.8438	8.885720e+05
licenciatura	524296.5000	5.007508e+05
matrimonios	10486.3438	8.013488e+03
nacimientos	50797.6250	3.793914e+04
victimasAcc	119.5625	8.613753e+01
pobreza	1739196.8750	1.695164e+06
religion	3323228.7812	2.829243e+06
afiliadosSalud	2893212.8750	2.220232e+06
transtornos	160.8406	1.210800e+02

Figure 2: Vector de Medias y Desviación Estándar de las Variables

$$\eta(\mu) = \beta_0 + \beta_1 X_1 + \dots + \beta_n X_n,$$

con $\mu = E(Y|X)$ y η como función liga.

6.2 Justificación de la selección

La selección de un Modelo Lineal Generalizado (GLM) se justifica por su versatilidad y capacidad para abordar diversas características de los datos demográficos. Los GLM ofrecen flexibilidad al modelar diferentes distribuciones de datos, lo que es fundamental al analizar variables de respuesta como las tasas de mortalidad y natalidad, que pueden no seguir una distribución normal. Además, los GLM permiten manejar relaciones no lineales entre variables predictoras y de respuesta, lo que es relevante en estudios demográficos donde estas relaciones pueden ser complejas. Su capacidad para considerar la dispersión y la estructura de correlación entre observaciones es esencial en el análisis de datos demográficos, donde tales características son comunes. Los coeficientes estimados en un GLM tienen interpretaciones directas, facilitando la comprensión de la relación entre las variables. Además, los GLM son aplicables a

una amplia gama de tipos de datos, incluyendo variables binarias, de conteo y continuas, lo que los convierte en una herramienta importante para analizar diversas facetas de las tasas de mortalidad y natalidad en la población.

Las desventajas del proceso de este modelo, consisten en la importancia de seleccionar eficazmente a tus variables, pues el modelo podría no aportar mucho al considerar tantas entradas; así mismo, elegir la distribución de la variable de respuesta, como la función liga, puede llegar a ser delicado, aunque existen métodos para seleccionar el mejor modelo.

6.3 Descripción Técnica

Se realizarán 2 ajustes de modelo, uno correspondiente a la natalidad y el otro para la mortalidad. Es decir, $Y_1 = \text{Natalidad}$, $Y_2 = \text{Mortalidad}$

El análisis se llevará a cabo utilizando software estadístico, como R. Primeramente, se utilizará una malla para jugar con los modelos continuos y con las funciones ligas que les corresponderían, después se considerará la métrica BIC para seleccionar entre todas las combinaciones de modelos el que tenga menor BIC; luego checaremos supuestos para ver si es apropiado. Al no observar que sea apropiado, utilizaremos alternativas de distribuciones discretas: Poisson y Binomial Negativo. Posteriormente, se utilizará optimización Lasso¹ para considerar que variables no aportan nada al modelo e ir las retirando; volveremos a checar supuestos de estos modelos modificados.

7 Resultados

7.1 Resultados Descriptivos

Las variables fueron reducidas en poca parte, pues de las 15 variables explicativas, solo resultaron significativas 11 en total para ambos modelos. Con ayuda de R probamos los diferentes modelos comparando cada BIC y checando los supuestos de cada uno; la optimización lasso fue fundamental, pues a partir de ella se hizo la primera reducción de variables. Ya posteriormente por criterio propio y para obtener la mayoría de variables significativas, se retiraron las demás variables. Obtuvimos que ambos modelos deberían de ser considerados la distribución Binomial Negativa (línea 504-521 y 700-720), término offset y función liga $\eta = \ln$.

7.2 Estimación del modelo

Con

$$E(Y_1|X) = \mu_1, \quad E(Y_2|X) = \mu_2,$$

tenemos que la estimación de los modelos resultó

¹Es una técnica de regresión que se utiliza tanto para la selección de variables como para la regularización, con el fin de mejorar la precisión predictiva y la interpretabilidad del modelo estadístico resultante. El objetivo es minimizar la suma de los errores al cuadrado entre las predicciones del modelo y los valores reales. La regresión Lasso introduce una penalización en esta minimización que es proporcional a la suma de los valores absolutos de los coeficientes del modelo.

$$\begin{aligned}\ln(\mu_1) = & \ln(\text{PoblacionTotal}) - 4.297364 \\ & - 1.885502 \times 10^{-7} \text{Alfabetismo} \\ & - 8.609721 \times 10^{-8} \text{Emigración} \\ & + 7.626962 \times 10^{-8} \text{Pobreza} \\ & + 1.470856 \times 10^{-7} \text{AfiliadosSalud}\end{aligned}$$

$$\begin{aligned}\ln(\mu_2) = & \ln(\text{PoblaciónTotal}) - 4.908649 \\ & + 6.742680 \times 10^{-7} \text{Alfabetismo} \\ & + 2.549774 \times 10^{-7} \text{Discapacidad} \\ & - 1.777477 \times 10^{-7} \text{Inmigración} \\ & - 5.386475 \times 10^{-7} \text{Licenciatura} \\ & + 7.626962 \times 10^{-8} \text{Pobreza} \\ & - 5.157851 \times 10^{-7} \text{Religión}\end{aligned}$$

8 Discusión

8.1 Análisis de Resultados

Para el modelo ajustado de nacimientos tenemos:

- El intercepto es de -4.297364. Esto implica que si todas las demás variables son 0, se esperaría que la tasa de nacimientos sea de $e^{-4.297364}$, lo que puede interpretarse como una tasa baja de nacimientos.
- Los coeficientes de las variables tienen las siguientes interpretaciones:
 - El coeficiente para "Alfabetismo" es -1.885502×10^{-7} . Esto sugiere que un aumento en el nivel de alfabetismo está asociado con una ligera disminución en la tasa de nacimientos, manteniendo constantes las demás variables del modelo.
 - El coeficiente para "Emigración" es -8.609721×10^{-8} . Indica que un aumento en la emigración está asociado con una ligera disminución en la tasa de nacimientos, manteniendo constantes las demás variables del modelo.
 - El coeficiente para "Pobreza" es 7.626962×10^{-8} . Esto sugiere que un aumento en la pobreza está asociado con un ligero aumento en la tasa de nacimientos, manteniendo constantes las demás variables del modelo.
 - El coeficiente para "AfiliadosSalud" es 1.470856×10^{-7} . Indica que un aumento en el número de personas afiliadas a algún servicio de salud está asociado con un ligero aumento en la tasa de nacimientos, manteniendo constantes las demás variables del modelo.
- La significancia de las variables se indica mediante el p-value. La variable que no resulta significativa es "AfiliadosSalud" (línea 497), lo cual se puede deber a que está altamente correlacionada con otra variable y no aporta más información al modelo; se decidió mantenerla porque al retirarla, las demás variables resultaban modificadas en su significancia.

- En términos de la magnitud de los efectos, los coeficientes más grandes corresponden a las variables "Alfabetismo" y "AfiliadosSalud", lo que sugiere que estas variables tienen un impacto relativamente más fuerte en la tasa de nacimientos en comparación con las otras variables del modelo.

Ahora, lo que observamos del modelo ajustado a las defunciones:

- El intercepto es de -4.908649. Esto implica que si todas las demás variables son 0, se esperaría que la tasa de defunciones sea de $e^{-4.908649}$, lo que puede interpretarse como una tasa baja de defunciones.
- Los coeficientes de las variables tienen las siguientes interpretaciones:
 - El coeficiente para "Alfabetismo" es 6.742680×10^{-7} . Esto sugiere que un aumento en el nivel de alfabetismo está asociado con un ligero aumento en la tasa de defunciones, manteniendo constantes las demás variables del modelo.
 - El coeficiente para "Discapacidad" es 2.549774×10^{-7} . Esto sugiere que un aumento en la discapacidad está asociado con un ligero aumento en la tasa de defunciones, manteniendo constantes las demás variables del modelo.
 - El coeficiente para "Inmigración" es -1.777477×10^{-7} . Indica que un aumento en la inmigración está asociado con una ligera disminución en la tasa de defunciones, manteniendo constantes las demás variables del modelo.
 - El coeficiente para "Licenciatura" es -5.386475×10^{-7} . Esto sugiere que un aumento en la población con licenciatura está asociado con una ligera disminución en la tasa de defunciones, manteniendo constantes las demás variables del modelo.
 - El coeficiente para "Pobreza" es 7.626962×10^{-8} . Indica que un aumento en la pobreza está asociado con un ligero aumento en la tasa de defunciones, manteniendo constantes las demás variables del modelo.
 - El coeficiente para "Religión" es -5.157851×10^{-7} . Esto sugiere que un aumento en la religión está asociado con una ligera disminución en la tasa de defunciones, manteniendo constantes las demás variables del modelo.
- La significancia de las variables se indica mediante el p-value y obtuvimos que todas de ellas (línea 683), a excepción de "Licenciatura", son significativas (0.05) en el modelo.
- En términos de la magnitud de los efectos, los coeficientes más grandes corresponden a las variables "Alfabetismo", "Licenciatura" y "Religión", lo que sugiere que estas variables tienen un impacto relativamente más fuerte en la tasa de defunciones en comparación con las otras variables del modelo.

8.2 Conclusión

Se pudieron encontrar factores demográficos significativos (alfabetismo, la afiliación a servicios de salud y la población que cuenta con licenciatura) que afectan las tasas de mortalidad y natalidad de México, así como el comportamiento esperado de las tasas al verse modificados estos factores. Cabe resaltar que nos limitamos a analizar estas tasas con factores demográficos, pues considero que puede haber otro tipo de variables que perjudicarían a las tasas como enfermedades, contaminación, consumo de alcohol o tabaco, estrato social, entre otras; éstas suenan bastante prometedoras, sin embargo, el hecho de ser estricto con las fuentes que proporcionan los datos, resultó en no incluirlas. A diferencia de estudios previos Chillagana en (2019), el

modelo ajustado resulta más compacto y simple a comparación de una serie de tiempo que se estimó en su estudio, así como el uso de otras variables como las actividades físicas, consumo de frutas y verduras y consumo de cigarrillos. Además, la autora parte de teoría económica y uso de modelos econométricos para explicar las tasas. Este estudio sienta algunas bases para futuras investigaciones que podrían explorar cómo estas tendencias demográficas influirían en diversos aspectos de la sociedad mexicana, como la economía, la salud pública y las políticas sociales; sumado a eso, se espera que en el futuro se tomen en cuenta otros tipos de factores, fuera de lo demográfico, para comprender mejor las tendencias y ofrecer una explicación más completa en este tema de interés. A partir de estos hallazgos, se pueden sintetizar las siguientes recomendaciones para responsables de políticas, profesionales de la salud y otros actores interesados en mejorar las tasas de mortalidad y natalidad en México:

- **Mejorar la Educación:**

Implementar programas de educación con la finalidad que mayor población que cuente con educación superior, lo cual está asociado con menores tasas de mortalidad.

- **Fortalecer los Servicios de Salud:**

Aumentar la afiliación a servicios de salud, ya que una mayor cobertura está relacionada con un aumento en las tasas de natalidad, posiblemente debido a un mejor acceso a cuidados prenatales y planificación familiar.

- **Atención a Poblaciones Vulnerables:**

Focalizar políticas de salud y sociales en regiones y segmentos de la población con altos niveles de pobreza y discapacidad, ya que estos factores están asociados con mayores tasas de mortalidad.

- **Incentivar la Inmigración:**

Facilitar procesos y condiciones que favorezcan la inmigración, dado que esta variable muestra una relación con la disminución de las tasas de mortalidad.

Sin embargo, es importante señalar que no se tuvo éxito en el objetivo de corroborar y analizar las tendencias y los patrones ya que, aunque se logró ajustar un modelo demográfico, no se cuentan con datos actuales para validar si las tasas estarían bien estimadas.

9 Referencias

- Tasa de mortalidad. examen de los factores que afectan la frecuencia de las muertes - FasterCapital. (s.f.). FasterCapital. <https://fastercapital.com/es/contenido/Tasa-de-mortalidad-examen-de-los-factores-que-afectan-la-frecuencia-de-las-muertes.html>
- Factores que afectan la tasa de natalidad - FasterCapital. (s. f.). FasterCapital. <https://fastercapital.com/es/tema/factores-que-afectan-la-tasa-de-natalidad.html>
- Chillagana, M. (2019). Búsqueda de los factores influyentes en las tasas de natalidad, fertilidad, mortalidad general e infantil en Europa. [Grado de administración y dirección de empresas]. Univeristat Politècnica de Valencia.