

Factores económicos y sanitarios que influyen en la esperanza de vida: análisis de datos internacionales (2013–2023).

1. Código

Para la realización de este proyecto, se utilizó el lenguaje de programación R, que es el que estamos viendo en el curso.

El código está dividido en 6 partes distintas entre sí: Carga de paquetes, Carga de df, gráficos de variables, gráficos de relaciones, Correlaciones y por último Estadísticas útiles; en cada apartado se desarrollan diferentes tareas las cuales van a ser explicadas a continuación.

1-Carga de paquetes:

Lo más básico antes de empezar un proyecto en R, cargar los paquetes que sean necesarios para el desarrollo del proyecto.

Utilicé **library()** para cargar los paquetes de tidyverse, este para diferentes funciones útiles en R, y **readr** para poder leer los .csv desde internet.

2-Carga de df

En este apartado definimos una serie de países (Unos generales y otros filtrados) esto para después filtrar dichos países en DataFrame que descarguemos. Con ayuda de **readr**, cargamos los csv desde la página de internet, no hace falta descargar ningún csv manualmente ya que la página nos proporciona la facilidad de importar el csv desde una URL.

Dentro de esta sección también modificamos la data para facilidad de elaboración, se utilizó **filter()** para filtrar los países (con el vector definido anteriormente) y los años para que coincidan con los años que se quieren estudiar. Con ayuda de una inteligencia artificial, se creó un **df_continentes**, el cual contiene el nombre del país en la primera columna, y en la segunda columna el continente al que pertenece.

Como se usaron diferentes bases de datos para este proyecto, decidí utilizar **left_join()**, para poder unir la información dependiendo del país, años y código de país. De aquí sale el nuevo df, llamado **df_final**, que es la base de datos sobre la cual gira el proyecto, además nos sirve para hacer la base de datos filtrada que es para hacer el estudio un poco mas focalizado en ciertos países.

Por último, para una lectura más fácil de la base de datos, se utilizó un **transmute()**, para ordenar las columnas en el orden deseado.

3-Graficos de Variables:

Acá se desarrollan los distintos gráficos de las variables de la base de datos, que muestran el comportamiento de dichas variables en el intervalo de tiempo elegido; como también se quiere hacer un estudio mas focalizado, en base al df_filtrado se hacen los mismos gráficos, solo que acá se muestran los países filtrados.

Tenga en cuenta que los gráficos generales no son tan fáciles de leer, debido a la gran cantidad de países que hay, pero nos da una idea de como va la tendencia. Los gráficos filtrados están justamente para detallar más cómodamente los países seleccionados.

4-Graficos Relacionados:

Esta sección es muy interesante, ya que acá se desarrollan los gráficos que muestran los gráficos donde podemos ver la relación entre dos variables. Como anteriormente se comentó, tenemos dos enfoques, uno mas general y el caso filtrado, en estos gráficos se utilizó el **geom_smooth()** para poder analizar más fácilmente la tendencia, ya sea la general, o en el caso filtrado, la tendencia por país

5-Correlaciones

Se utiliza la función **cor(.. , method = “Pearson”)** para obtener el coeficiente de correlación entre dos variables.

6-Estadísticas Útiles

En estadísticas útiles, lo que se desea hacer es como una especie de rankings dependiendo de la variable, se utilizaron funciones como **filter() summarise() arrange() group_by()**, las cuales se vieron en la primera mitad del curso y son muy útiles a la hora de hacer este tipo de rankings

2. Resultados

Esta parte de resultados vamos a dividirla en varias partes:

1-Analisis descriptivo por Variable

- Esperanza de vida: En el caso general, se puede observar que, en la mayoría de los países, conforme aumentan los años, la esperanza de vida aumenta, claramente hay una caída por el año 2020, donde sucedió la pandemia por COVID-19. Para el caso filtrado, es lo que se observó en el grafico general, solo que algunos países parecen estar estancados, mientras que otros si aumentan considerablemente.
- Gasto en salud per cápita: Para el grafico general, podemos notar la diferencia que hay entre los que gastan en promedio algunos países, mientras que unos gastan mas de \$10000, otros no llegan ni a los \$1000. Viendo el grafico filtrado, la diferencia se nota aún más.
- Tasa de mortalidad infantil: Se puede evidenciar una leve tendencia a la baja para la mayoría de los países, podemos observar que hay países con tasas de mortalidad muy bajas, mientras que otros países tienen tasas bastante elevadas.
- PIB per cápita: Una vez más podemos observar la diferencia que hay entre los países, hablando estrictamente de lo económico, mientras hay países con PIB per cápita muy elevados, países mas pobres, tiene valores de PIB extremadamente bajos.

2-Graficos Relacionados:

- Gasto en salud vs Esperanza de vida:
En el grafico general, podemos observar una relación positiva, entre más se gasta en salud, mas es la esperanza de vida; pero esto no es para todos los países, en el caso filtrado podemos observar que hay países con relaciones negativas, gastan mucho y no aumenta la esperanza de vida, y se destaca el caso de Costa Rica, que, con un gasto no tan elevado, tiene una esperanza de vida alta.
- PIB per cápita vs Esperanza:
Podemos observar que la tendencia general nos dice que países mas ricos, mayor esperanza de vida, esto se puede deber a una mejor calidad de vida, la cual colabora con que la esperanza de vida aumente, hay más recursos para las personas; aunque esta relación no es muy perfecta, porque hay una gran aglomeración en el nivel de esperanza de vida entre los 80 y 85 años.

- PIB per cápita vs Mortalidad infantil:
Vemos que conforme aumenta el PIB per cápita, la tasa de mortalidad infantil disminuye, esto también puede ir relacionado con la escasez de recursos, si el PIB es muy bajo, no hay muchas posibilidades de que los infantes puedan sobrevivir, esto se puede observar en países como Indonesia, que conforme más aumenta el PIB per cápita, menores son sus tasas de mortalidad
- Gasto en salud vs Mortalidad Infantil: En líneas más generales, podemos observar que entre más gastan las personas en salud, menor es la mortalidad infantil, esto porque si se gasta mas en salud, puede significar un mejor seguro de vida, el cual puede cubrir embarazos y medicamentos para los niños.

3-Análisis comparativo de los países seleccionados:

- Costa Rica:
Es un país que destaca por su alta esperanza de vida, a pesar de no ser un país de primer mundo, ni tener los mayores ingresos por persona en promedio. A pesar de no ser de los países que mas gastan en el mundo, mantiene una esperanza de vida alta, la cual habla de la eficiencia del sistema de salud de Costa Rica.
- Estados Unidos: Es uno de los países que más gastan, pero con resultados no tan eficientes, puede ser debido a los altos costos de la salud en este país. También tiene una esperanza de vida relativamente baja, esto puede ser debido a que Estados Unidos es un país en el cual constantemente hay problemas entre sus habitantes, y la facilidad de obtener armas es un factor importante.
- Japón: Como bien se sabe, Japón es un país donde la esperanza de vida es elevada, y aunque ellos no gasten tanto en salud, puede deberse a su estilo de vida y también a su sistema de salud, parece ser un sistema bastante eficiente. Además, podemos observar que la tasa de mortalidad es muy baja, una vez mas hacemos referencia al sistema de salud japones.
- Suiza: Al ser un país muy rico, tiene niveles muy buenos de esperanza de vida, aunque la relación gasto-esperanza, es una relación inversa, esto puede deberse al modelo de sistema sanitario que tienen, un sistema sanitario, que puede ser que se haya vuelto más caro, pero no tan eficiente.
- Indonesia: Al ser un país en desarrollo, comienza con niveles en casi todas las variables, pero con un efecto mas fuerte en el aumento de diferentes variables, por

ejemplo, entre más gastan en salud por persona, la mortalidad infantil y la esperanza de vida tienden a aumentar.

3-Limitaciones del análisis:

Hay que tener en cuenta que no se están tomando en cuenta factores como el estilo de vida ni el nivel de educación, solo se están considerando los expuestos en el proyecto.

4- Correlaciones entre variables

- | | | |
|--|-------|-------------------|
| ➤ Correlación | Gasto | Mortalidad |
| Se obtuvo un índice de correlación de -0.5013195, el cual nos indica que hay una relación negativa moderada, es decir que cuando una aumenta, la otra disminuye, pero no tan drásticamente | | |
| ➤ Correlación | PIB | Mortalidad |
| Se obtuvo un índice de correlación de -0.5475498, el cual nos indica que hay una relación negativa moderada, es decir que cuando una aumenta, la otra disminuye, pero no tan drásticamente | | |
| ➤ Correlación | PIB | Esperanza de Vida |
| Se obtuvo un índice de correlación de 0.5709529, el cual nos indica que hay una relación positiva moderada, cuando una aumenta, la otra aumenta, pero moderadamente | | |
| ➤ Correlación | Gasto | Esperanza |
| Se obtuvo un índice de correlación de 0.5915271, el cual nos indica que hay una relación positiva moderada, cuando una aumenta, la otra aumenta, pero moderadamente. | | |

3. Recomendaciones

Como recomendación para futuros análisis similares, se pueden tomar en cuenta otro tipo de variables para hacer el estudio más completo, por ejemplo, niveles de educación, acceso a servicios de salud, estilo de vida, enfermedades (obesidad, cáncer, enfermedades cardíacas, diabetes), pobreza, etc...

Se puede ampliar el tiempo estudiado, así se toman en cuenta eventos pasados, ya sean locales o globales. Además del tiempo se puede ampliar el estudio a diferentes zonas y/o regiones, ya que son aspectos importantes para hacer un buen análisis.

Jeremy Garcia Solano C33171

CA-0204

Herramientas para Ciencia de Datos I

Buscar esperanza de vida por género, ya que son muy diferentes entre sí, y pueden llegar a influir mucho en el estudio realizado.

4. Referencias

Our World in Data. (2025). *Child mortality – complete dataset [Conjunto de datos]*. Global Change Data Lab.

<https://ourworldindata.org/grapher/childmortality.csv?v=1&csvType=full&useColumnShortNames=true>

Our World in Data. (2025). *GDP per capita (World Bank dataset) [Conjunto de datos]*. Global Change Data Lab.<https://ourworldindata.org/grapher/gdp-per-capita-worldbank.csv?v=1&csvType=full&useColumnShortNames=true>

Our World in Data. (2025). *Health expenditure and financing per capita [Conjunto de datos]*. Global Change Data Lab.<https://ourworldindata.org/grapher/health-expenditure-and-financing-per-capita.csv?v=1&csvType=full&useColumnShortNames=true>

Our World in Data. (2025). *Life expectancy – complete dataset [Conjunto de datos]*. Global Change Data Lab.<https://ourworldindata.org/grapher/life-expectancy.csv?v=1&csvType=full&useColumnShortNames=true>