

PROYECTO FINAL

ELEMENTOS DE APRENDIZAJE DE MÁQUINA Y BIG DATA

Ausili Julián, Vanadía Matías

1. En Busca del Bienestar:

Explorando la Importancia de la Libertad Económica

2. Introducción

2.1. Contexto del Problema:

En un mundo interconectado, es crucial comprender las relaciones entre la libertad económica, la educación y la salud en diferentes países. Estos factores no solo impactan el desarrollo socioeconómico, sino que también influyen en la calidad de vida de las personas. Ante la creciente disponibilidad de datos a nivel global, se ha generado una necesidad de emplear técnicas avanzadas como el aprendizaje automático (Machine Learning) y el procesamiento de grandes volúmenes de datos (Big Data) para comprender y abordar problemas complejos.

En este proyecto, exploramos un diverso conjunto de indicadores que incluyen libertad económica, felicidad, esperanza de vida, PIB per cápita, tasas de suicidio y educación, con el objetivo de identificar patrones y agrupamientos significativos.

2.2. Importancia y Relevancia:

El estudio busca revelar patrones y conexiones entre los datos analizados. Esto permitirá entender cómo estas variables se entrelazan y afectan el bienestar de las poblaciones, proporcionando información valiosa para la toma de decisiones a nivel gubernamental y social.

La aplicación de técnicas de Machine Learning y Big Data permite una visión más profunda y sistémica, lo que puede llevar a la identificación de tendencias y la formulación de políticas más informadas.

2.3. Objetivos del Proyecto:

Este proyecto busca entender, mediante la aplicación de técnicas de Machine Learning y Big Data, cómo factores como la libertad económica, la felicidad y otros influyen en el bienestar de las personas en diferentes países.

Específicamente, buscamos identificar agrupamientos que no solo reflejen similitudes en los datos, sino también proporcionar información valiosa sobre qué aspectos específicos están vinculados a una mejor calidad de vida.

Además, al comprender cómo estos factores se relacionan entre sí y con el bienestar general, se pueden diseñar políticas más efectivas que mejoren la calidad de vida de las personas en todo el mundo.

3. Metodología

3.1. Datos Utilizados:

Para llevar a cabo este análisis, se utilizaron diversas fuentes de datos, incluyendo indicadores de libertad económica, felicidad, esperanza de vida, PIB per cápita, tasas de suicidio y datos educativos a nivel mundial. Estos conjuntos de datos se integraron y limpiaron para garantizar la coherencia y la calidad de los resultados finales. anexo 1

3.2. Herramientas y Tecnologías:

El proyecto se implementó utilizando Google Colab, aprovechando las bibliotecas de Python como pandas, geopandas, numpy, matplotlib y plotly para la manipulación, visualización y análisis de datos. Se emplearon algoritmos de clustering K-Means (sklearn) para identificar patrones en los datos. anexo 2

3.3. Proceso de Análisis/Desarrollo:

Antes de abordar el análisis, se realizó una exhaustiva búsqueda de conjuntos de datos relevantes para garantizar una cobertura integral de las variables de interés. La selección final incluyó datasets que abarcaban desde indicadores económicos y sociales hasta aspectos de bienestar.

Es importante destacar que algunos países se representan en color blanco en los mapas debido a la falta de datos disponibles en el conjunto de datos utilizado. La ausencia de información para estos países puede deberse a diversas razones, como la falta de informes oficiales, restricciones de acceso a datos, o limitaciones en la recopilación de información. En consecuencia, se recomienda interpretar los espacios en blanco como áreas donde la información correspondiente no está disponible y, por lo tanto, no se puede evaluar la situación específica en esos lugares. Esta consideración es fundamental para una interpretación precisa de los resultados y para evitar conclusiones incorrectas sobre las características de dichos países.

La metodología adoptada abarcó desde la carga y depuración inicial de datos hasta la aplicación de algoritmos de clustering. El proceso se dividió en varias etapas clave:

- Inicialmente, se estructuró el análisis en fases de limpieza de datos, exploración y visualización, así como la importación de las librerías necesarias.
- Se desarrollaron funciones especializadas para la generación de diversos gráficos interactivos, incluyendo barras, líneas, puntos, vistas geoespaciales y proyecciones en globo terráqueo.

- Se llevaron a cabo diversas visualizaciones interactivas para cada indicador, permitiendo comparaciones a nivel mundial.
- Además, se implementaron técnicas de clustering para agrupar países, explorando diferentes números de clusters con el fin de identificar la estructura subyacente más apropiada en los datos. Este enfoque proporcionó una comprensión más profunda de las relaciones y similitudes entre las naciones, enriqueciendo la interpretación de los resultados.

4. Resultados

4.1. Presentación de Resultados:

- Libertad económica:

Se han creado mapas interactivos que visualizan la libertad económica en diferentes países, así como gráficos de barras que muestran los países con la mayor y menor libertad económica. Estos gráficos proporcionan una visión general de cómo se distribuye este índice a nivel mundial.

4.1.1. En el mapa global se observa una variabilidad sustancial en los niveles de libertad económica en todo el mundo. Los tonos más claros indican mayores niveles de libertad, mientras que los más oscuros señalan restricciones en este aspecto clave para el desarrollo económico.

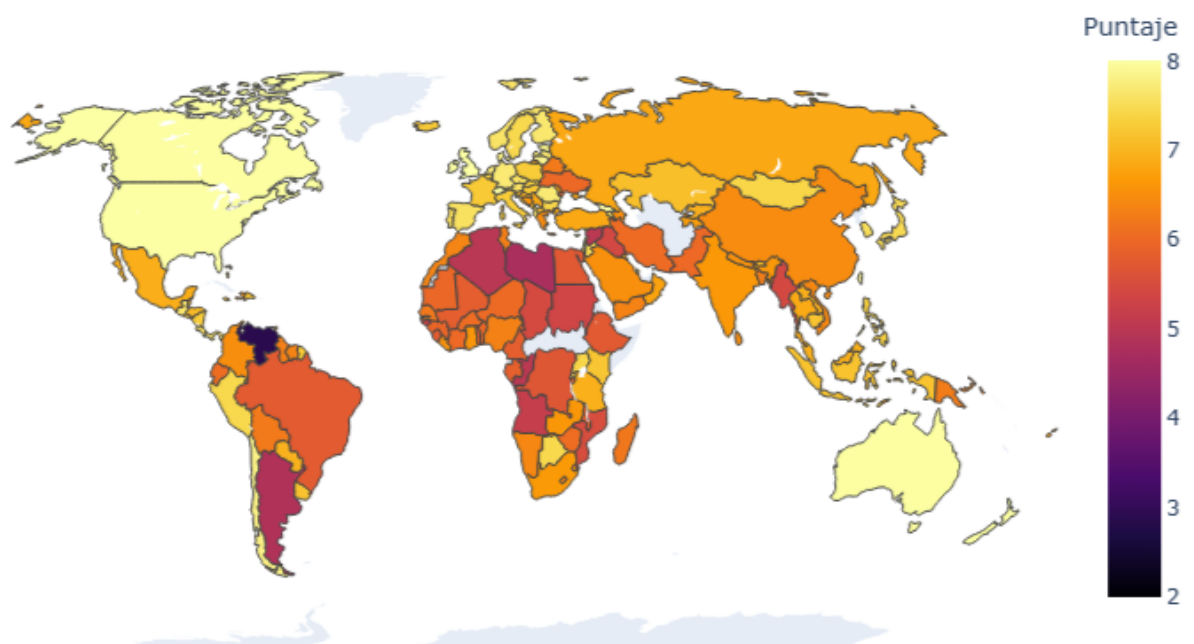


Figura 1.1: Mapa Global de Libertad Económica

4.1.2.

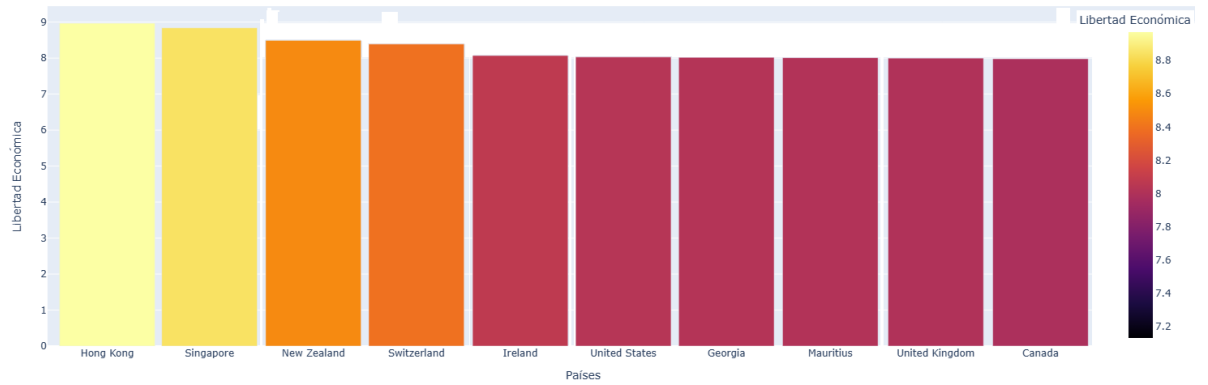


Figura 1.2.1: Top 10 Mejores Países en Libertad Económica

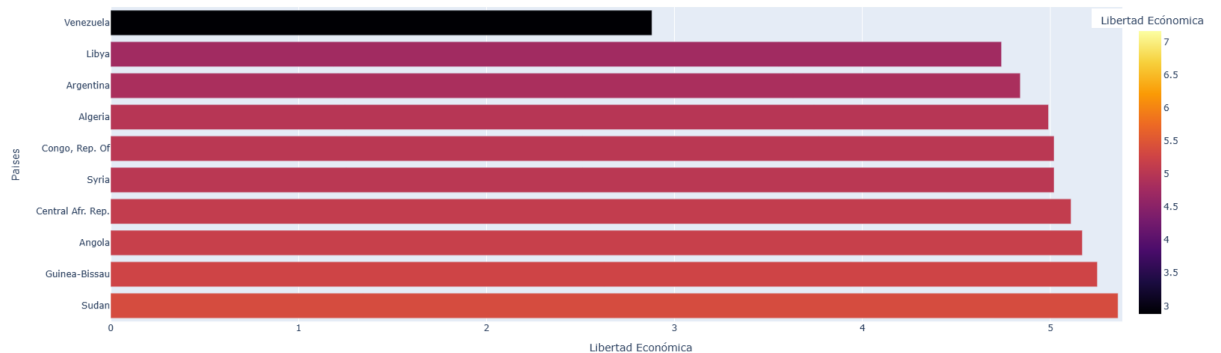


Figura 1.2.2: Top 10 Peores Países en Libertad Económica

- **Felicidad:**

Se ha generado un mapa de felicidad y un gráfico de barras que ilustra las diferencias en la felicidad entre los países.

4.1.3. El mapa proporciona una visión global de los niveles de felicidad en diferentes países. La escala de colores resalta patrones interesantes, donde los tonos azules indican puntajes más altos en este índice, mientras que los tonos rojizos representan puntajes más bajos. Esta representación permite explorar las variaciones geográficas en la satisfacción general de las poblaciones.

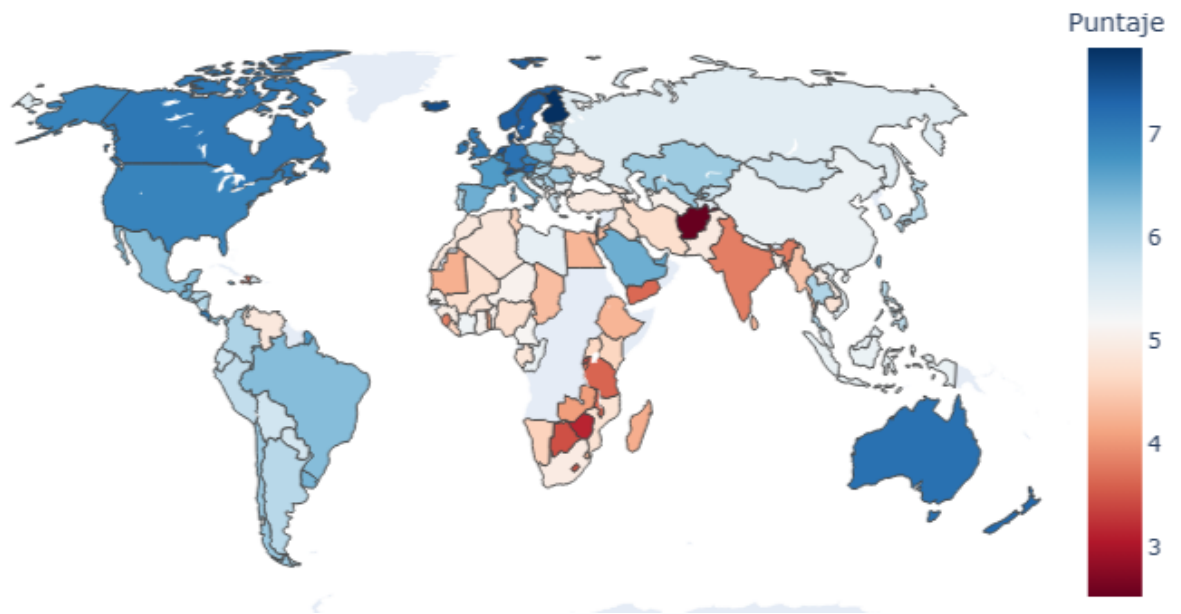


Figura 1.3: Mapa Global de Felicidad

4.1.4. El gráfico de barras compara los cinco países más felices con los cinco menos felices, proporcionando una visión clara de las diferencias en los niveles de felicidad

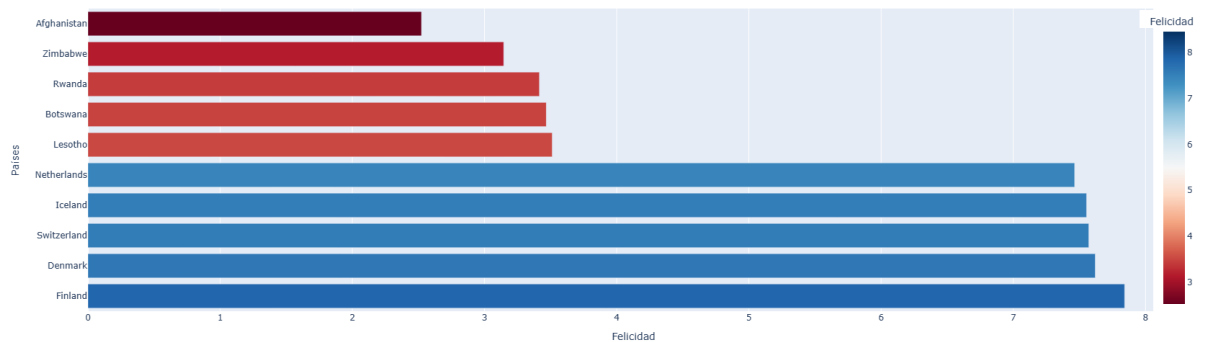


Figura 1.4: Top 5 Mejores y Peores Países en Felicidad

- **Esperanza de vida:**

Se presenta un mapa de esperanza de vida y un gráfico de barras que destacan las variaciones en la esperanza de vida. Se han identificado los países con las mayores tasas de esperanza de vida.

4.1.5. El mapa de esperanza de vida revela las diferencias en la expectativa de vida en diversas regiones del mundo. Los tonos más azulados representan niveles más bajos, mientras que los tonos verdes y amarillos indican expectativas de vida más altas.

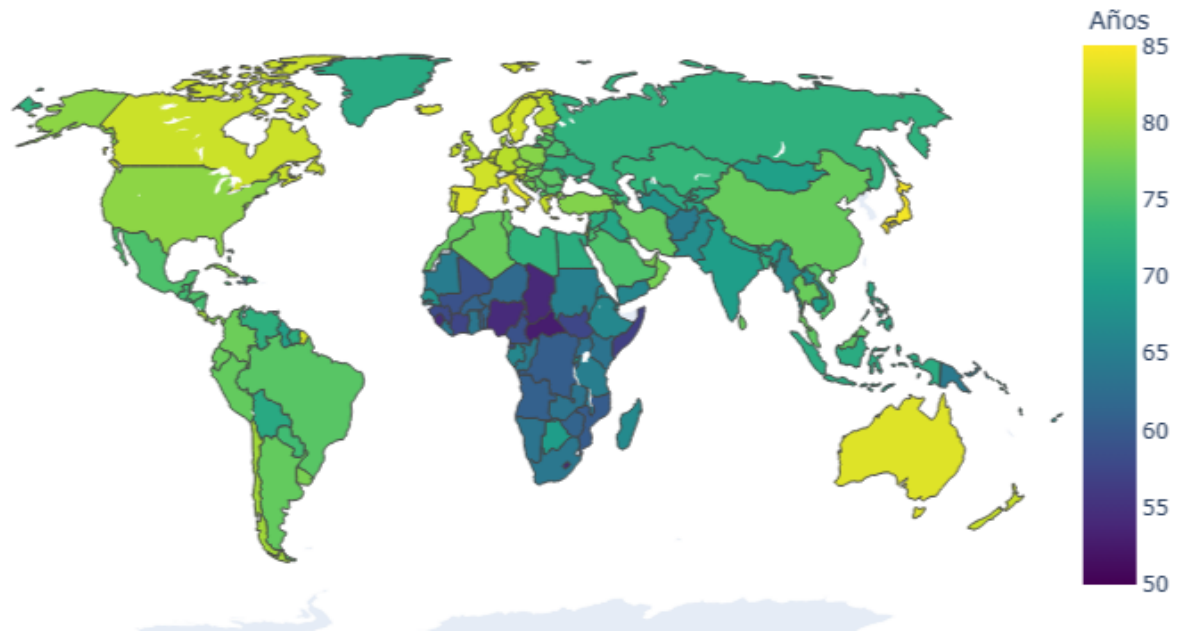


Figura 1.5: Mapa Global de Esperanza de Vida

4.1.6. El gráfico de barras destaca los 25 países con las expectativas de vida más altas. Entre los líderes en este indicador crucial para evaluar la calidad de vida se encuentran Hong Kong, Japón, Singapur, Italia y España, entre otros. Estos países sobresalen por sus altos estándares en la longevidad de sus poblaciones.

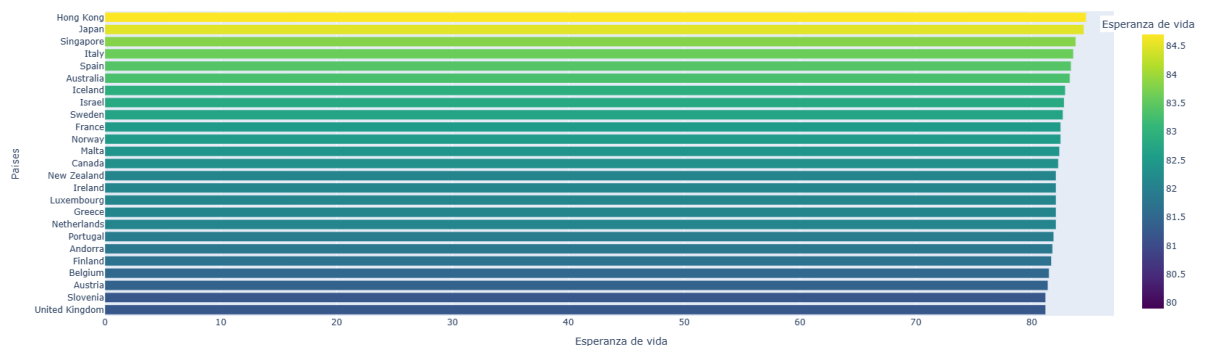


Figura 1.6: Top 25 Mejores Países en Esperanza de Vida

- **PBI per cápita:**

El PBI per cápita se ha visualizado en un mapa interactivo y en un gráfico de barras que muestra los países con los mayores y menores ingresos per cápita. Se han destacado las diferencias económicas globales.

4.1.7. El mapa de PBI per cápita ilustra las disparidades económicas en todo el mundo mediante una escala de colores que refleja los niveles relativos de ingresos por persona. Los tonos más oscuros indican niveles más bajos, mientras que los tonos

más claros representan ingresos per cápita más altos. Esta representación visual facilita la identificación rápida de las diferencias económicas globales.

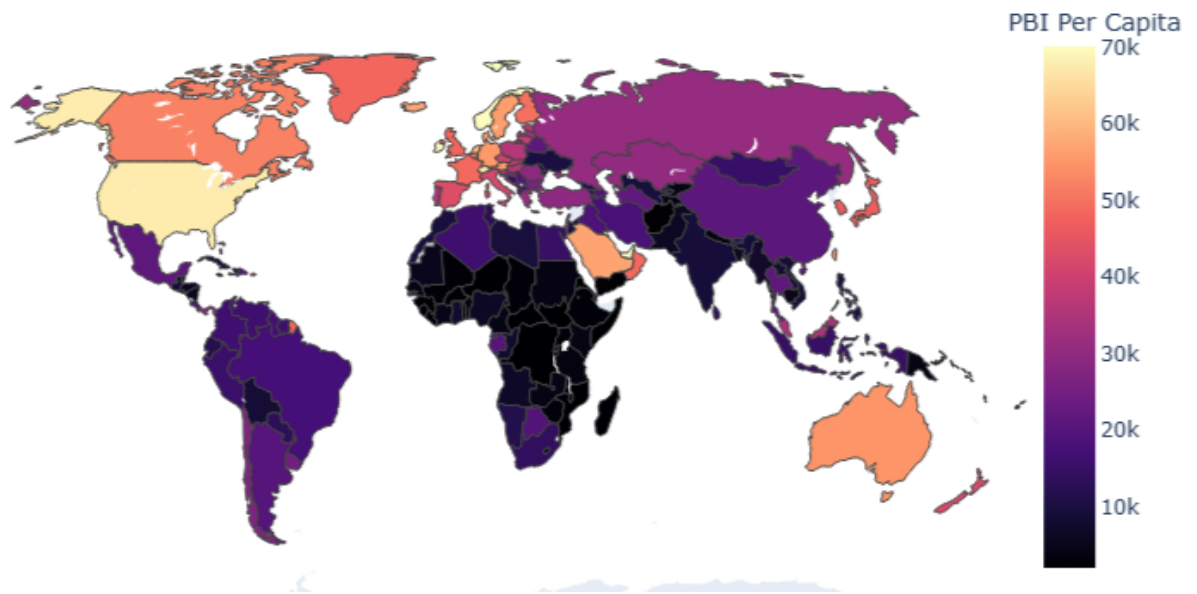


Figura 1.7: Mapa Global de PBI per Cápita

4.1.8. El gráfico de barras compara los diez países con el PBI per cápita más alto y más bajo. En la cima de esta métrica económica se encuentran Macau, Luxemburgo y Singapur, destacando por sus notables niveles de ingresos. En contraste, países como Somalia, Burundi y la República Centroafricana se ubican entre los de menor PBI per cápita, revelando disparidades económicas significativas. Este análisis proporciona una visión clara de las diferencias económicas extremas a nivel global.

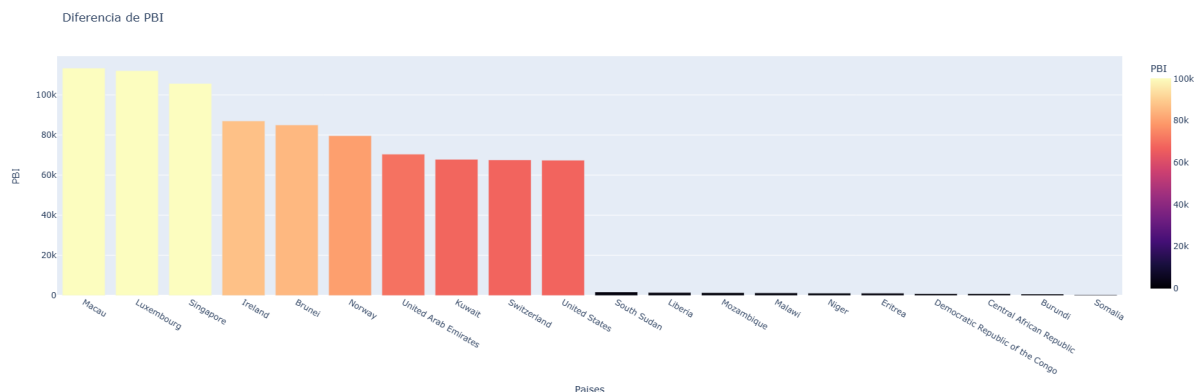


Figura 1.8: Top 10 Mejores y Peores Países en PBI per Cápita

- Tasa de suicidios:

Se ha explorado la tasa de suicidios en distintos países mediante un mapa temático. Se han identificado los países con las tasas más altas y más bajas de suicidios.

4.1.9. El mapa del ratio de suicidios ilustra las variaciones en las tasas de suicidio en diferentes regiones del mundo. La escala de colores resalta las diferencias en este indicador sensible, donde tonos más claros indican tasas más bajas y tonos más oscuros reflejan tasas más altas. Este análisis permite identificar patrones y áreas de preocupación en relación con la salud mental a nivel mundial.

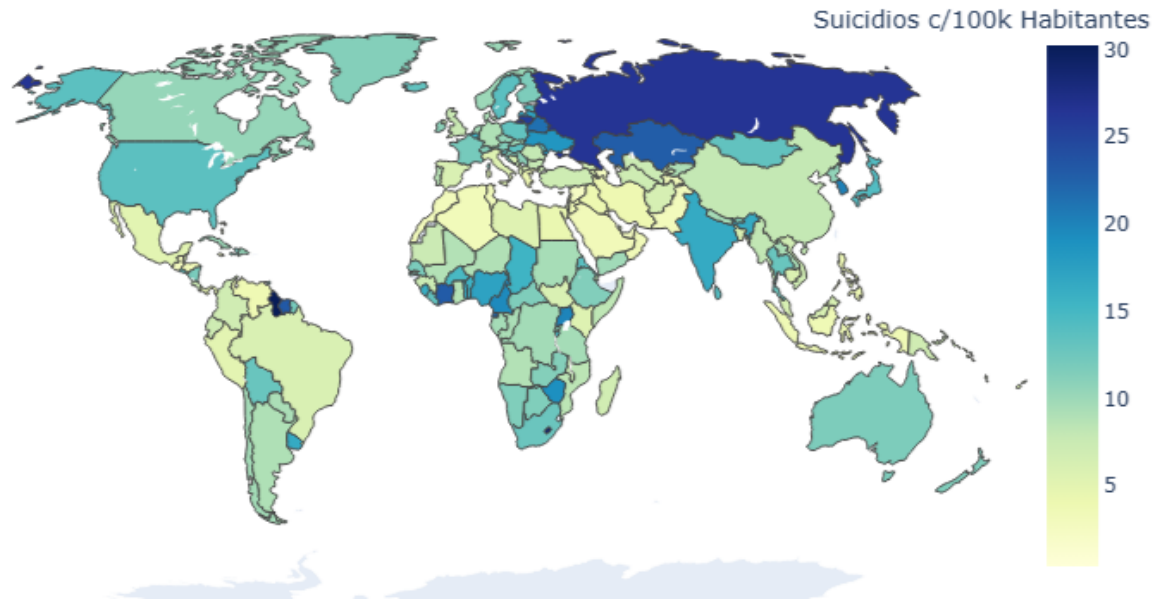


Figura 1.9: Mapa Global de Ratio de Suicidios

- Educación global:

Se han creado gráficos que abordan las tasas de no escolarización y finalización de la educación en distintos niveles. Estos gráficos ofrecen información sobre la accesibilidad y la calidad de la educación en diferentes países.

Países en blanco indican falta de datos en el conjunto utilizado. Interpretélos como áreas sin información disponible para evitar malentendidos sobre la educación en esos lugares.

4.1.10. Los mapas de no escolarización por niveles (preescolar, primaria, primer ciclo secundaria, segundo ciclo secundaria) ofrecen una visión detallada de la educación a nivel mundial. Los tonos rojos indican áreas con mayores desafíos en términos de no escolarización, mientras que los tonos más claros representan regiones con mejores indicadores educativos.

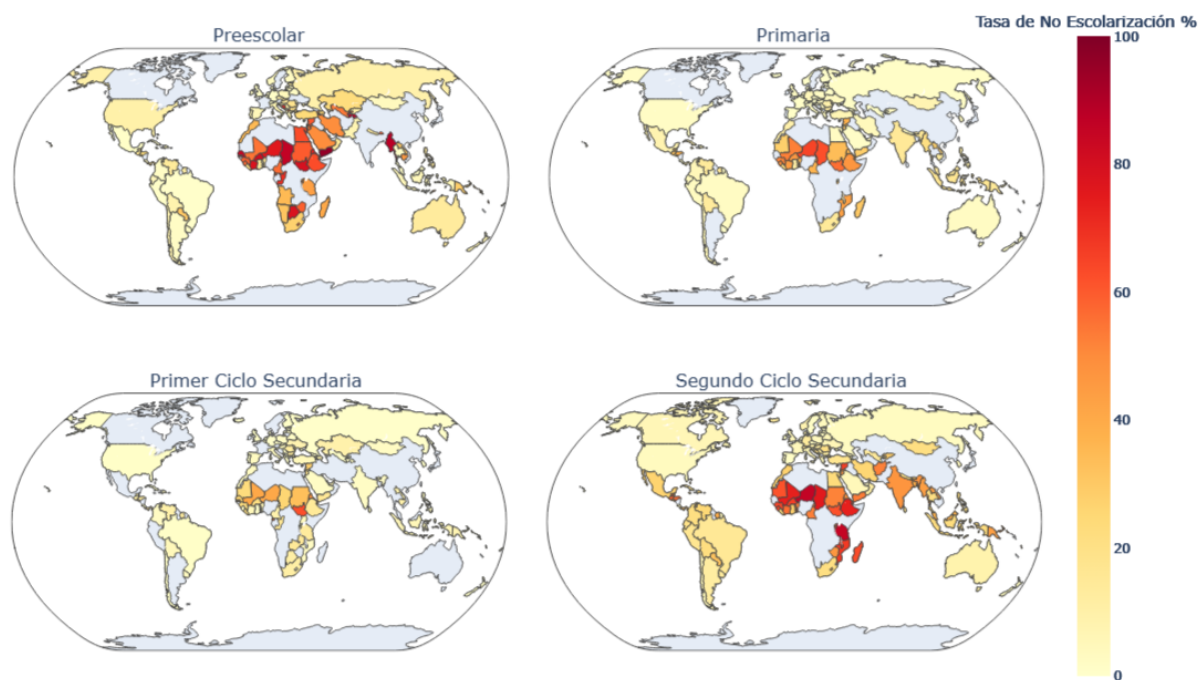


Figura 1.10: Mapa Global de No Escolarización por Niveles

4.1.11. Los mapas de finalización por niveles (Educación Primaria, Primer Ciclo Secundaria, Segundo Ciclo Secundaria) proporcionan información detallada sobre la conclusión de diferentes etapas educativas en todo el mundo. En este caso, tonos de verde más oscuro indican una mayor tasa de finalización educativa, destacando áreas donde se ha logrado un progreso significativo en términos de educación.

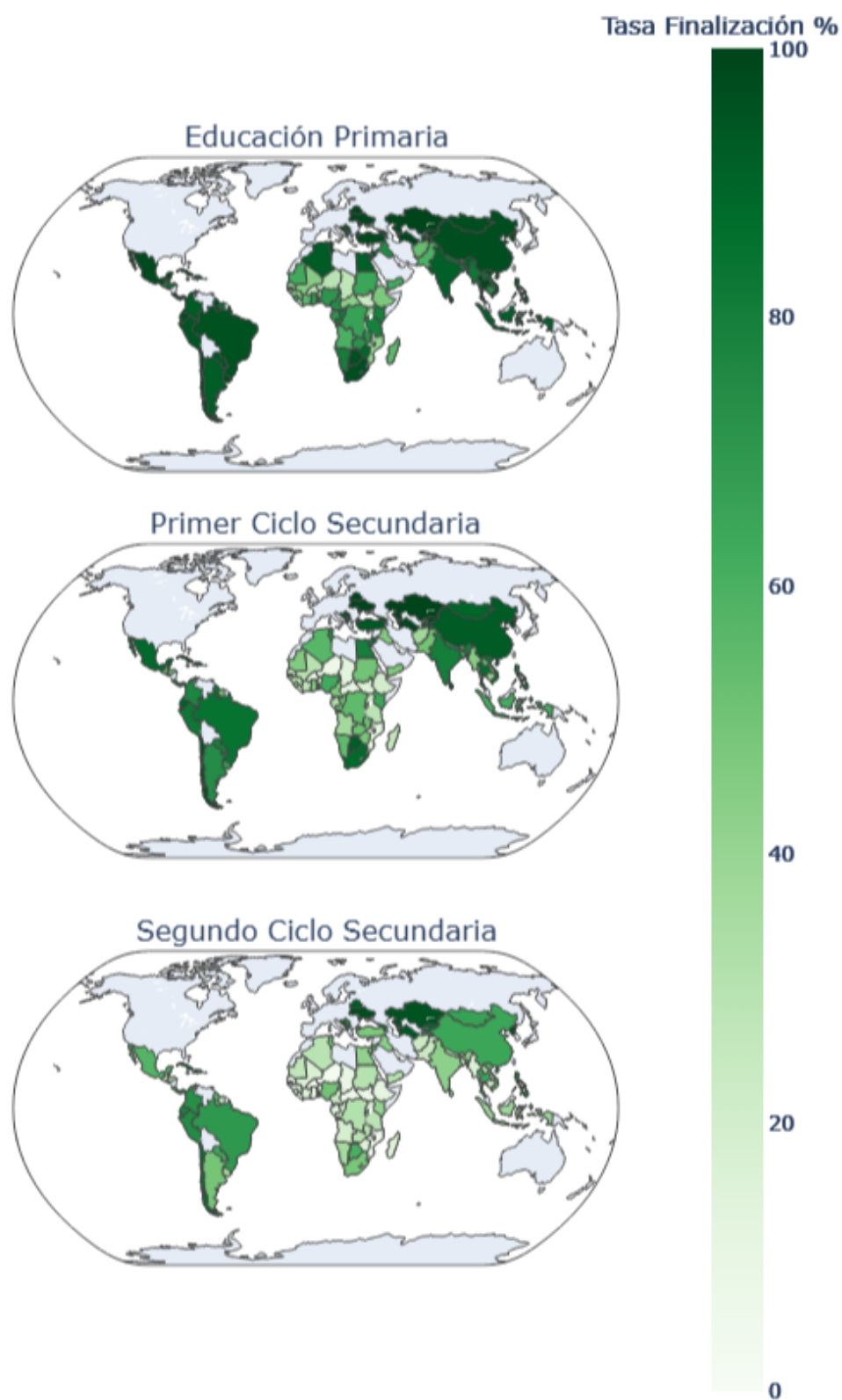


Figura 1.11: Mapa Global de Finalización por Niveles

- Desempleo:

4.1.12. El mapa de desocupación destaca las disparidades en las tasas de desempleo a nivel mundial. La escala de colores en tonos rojos ofrece una representación visual de las variaciones en este indicador clave, permitiendo identificar regiones con mayores desafíos en el ámbito laboral.

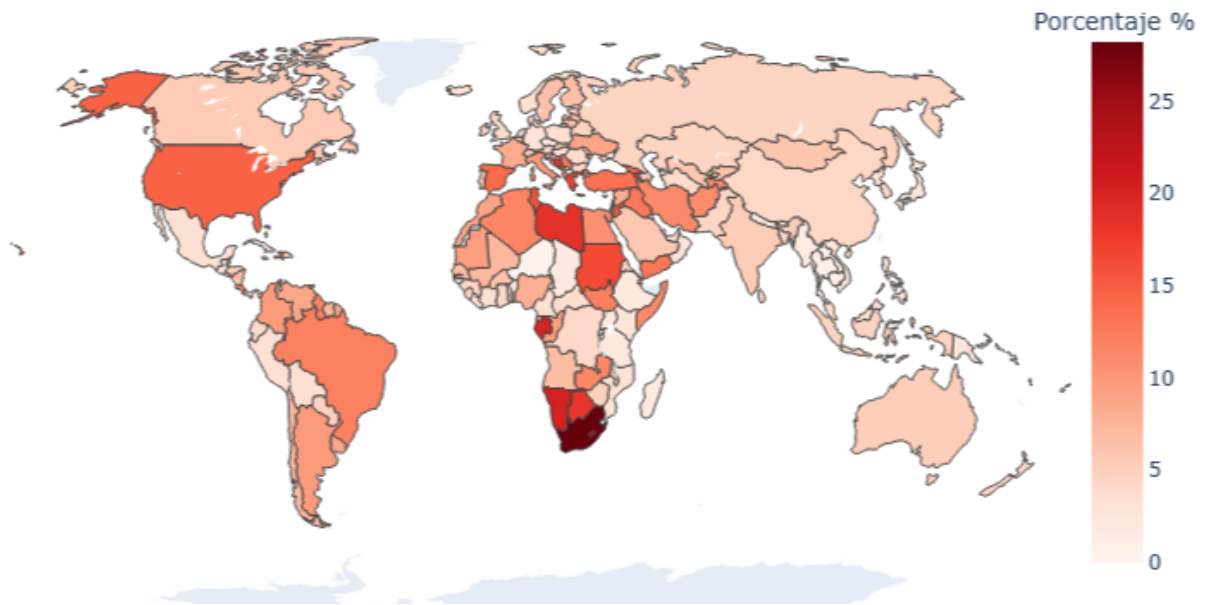


Figura 1.12: Mapa Global de Desocupación

- Clustering y calidad de vida:

Con el objetivo de identificar patrones más complejos, se ha aplicado el algoritmo k-means para agrupar países en función de múltiples indicadores. Se ha explorado el número óptimo de clusters, se han presentado mapas y gráficos que revelan la distribución de países en estos agrupamientos.

4.1.13. El gráfico de clusterización ayuda a determinar el número óptimo de clusters, proporcionando información valiosa para comprender las relaciones entre los indicadores clave (Libertad Económica, Felicidad, Tasa de Desempleo, PBI, Esperanza de vida, Ratio de suicidios).

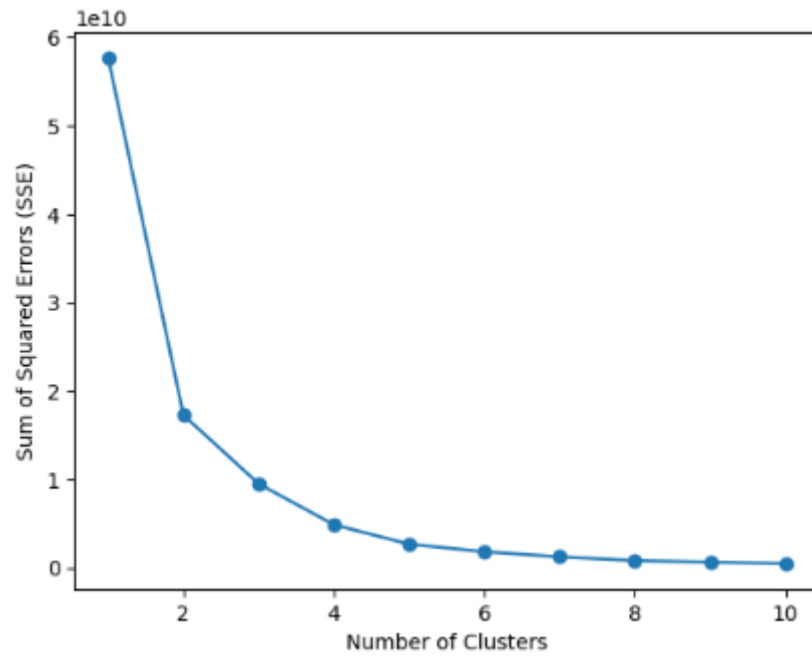


Figura 1.13: Gráfico de Clusterización para Encontrar el K Óptimo

4.1.14. El mapa de calidad de vida con tres clusters destaca las diferentes categorías de países, ofreciendo una visión general de las tendencias globales en la calidad de vida. Cada cluster representa un grupo de naciones con características socioeconómicas similares, brindando información sobre las disparidades y similitudes en términos de bienestar.

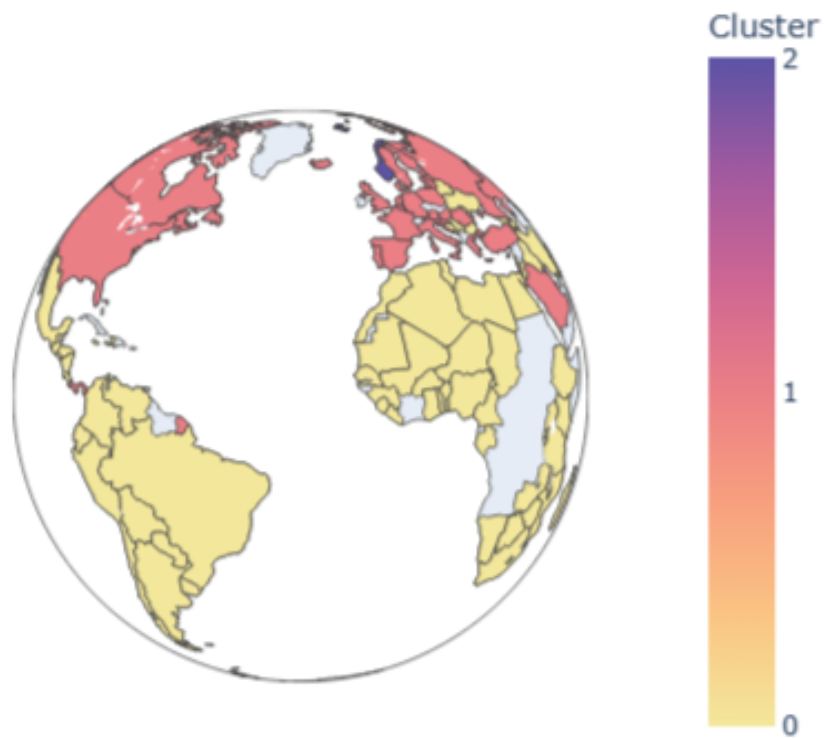


Figura 1.14: Mapa Global de Calidad de Vida (3 Clusters)

4.1.15. El gráfico de torta interactivo ilustra la distribución proporcional de los países en cada uno de los tres clusters, proporcionando una representación visual de la agrupación de naciones según sus características socioeconómicas y de calidad de vida.

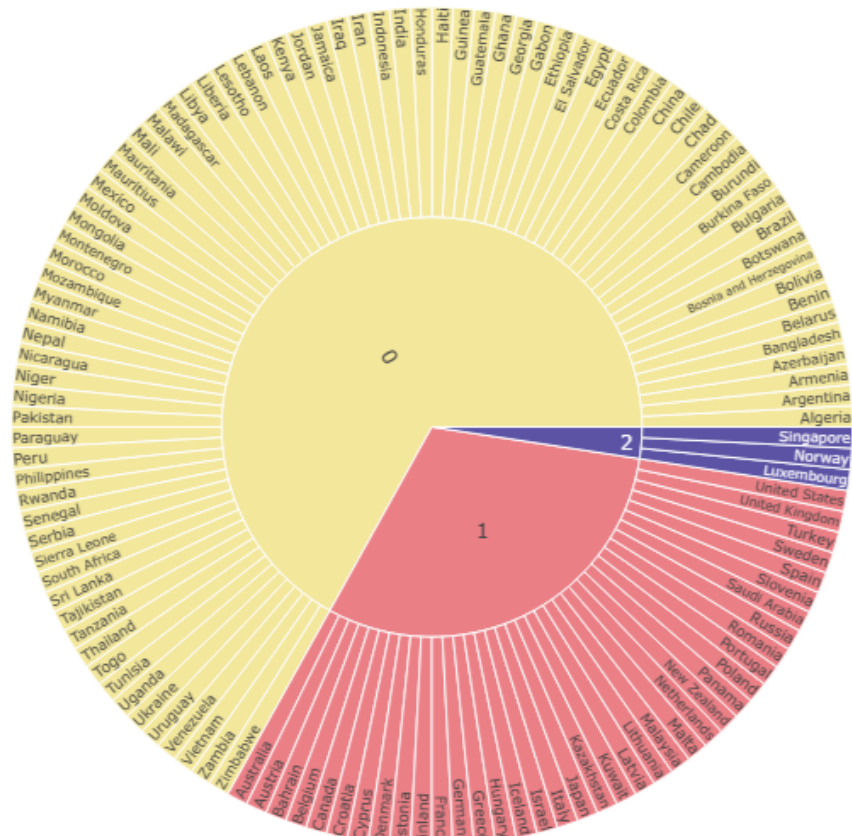


Figura 1.15: Gráfico Torta de Distribución de Países (3 Clusters) en Calidad de Vida

4.1.16. Al expandir el análisis a cuatro clusters, el mapa de calidad de vida proporciona una perspectiva aún más detallada de las variaciones en la calidad de vida a nivel mundial. Cada cluster representa un conjunto específico de países, resaltando las distintas categorías de bienestar. Esta mayor granularidad permite una comprensión más profunda de las dinámicas socioeconómicas y su impacto en la calidad de vida.

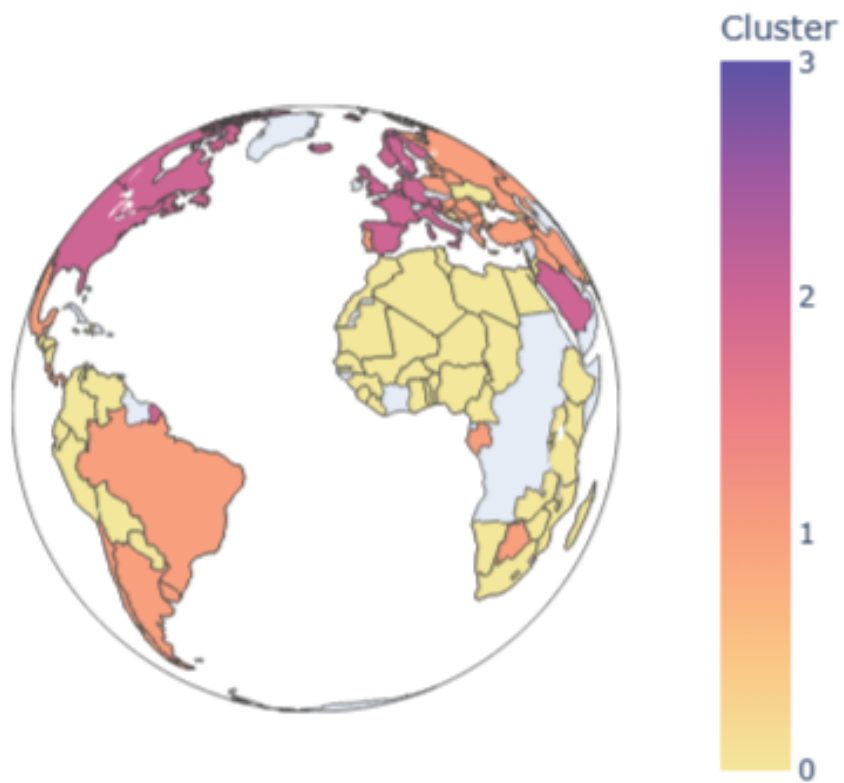


Figura 1.16: Mapa Global de Calidad de Vida (4 Clusters)

4.1.17. De manera similar, el gráfico de torta para el análisis de cuatro clusters muestra la distribución proporcional de los países en cada categoría, ofreciendo una visión más detallada de cómo se agrupan las naciones según sus características socioeconómicas y de calidad de vida.

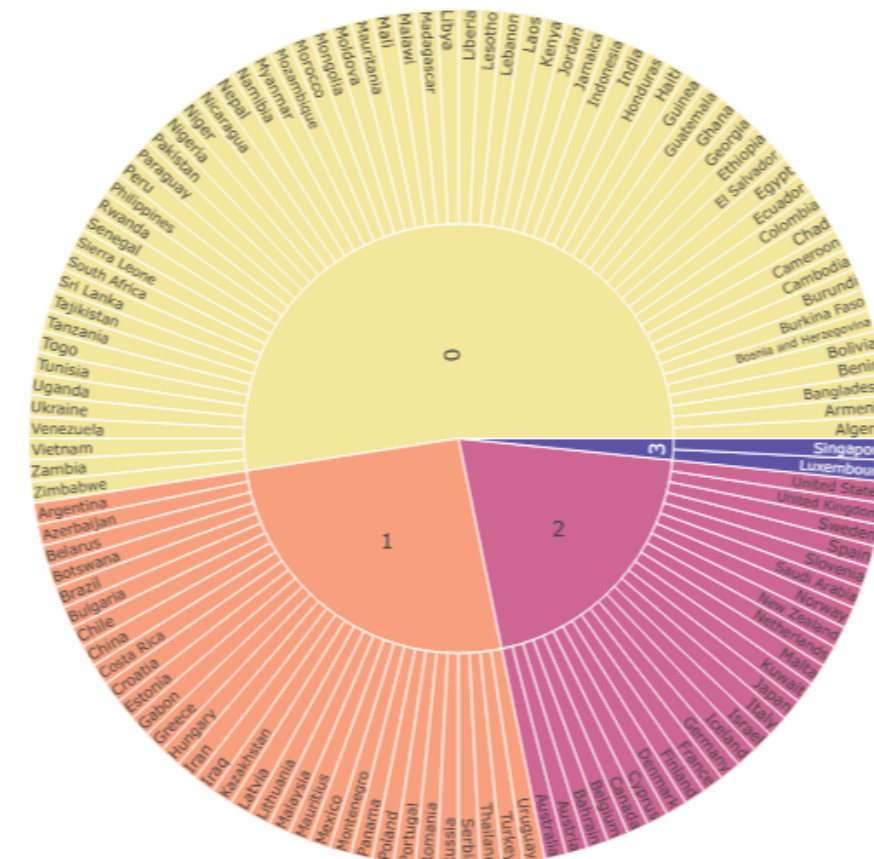


Figura 1.17: Gráfico Torta de Distribución de Países (4 Clusters) en Calidad de Vida

4.1.18. La matriz de correlación examina las relaciones estadísticas entre diferentes variables en nuestro conjunto de datos, permitiéndonos identificar patrones y asociaciones significativas.

Este análisis nos brinda una comprensión más profunda de cómo las diversas métricas, como la libertad económica, la felicidad y la esperanza de vida, están interrelacionadas en nuestro estudio integral de calidad de vida en el siglo XXI.



Figura 1.18: Matriz de Correlación

4.2. Interpretación:

4.2.1. Libertad Económica

Explorando los Matices del Desarrollo Global

El análisis del mapa global de libertad económica (figura 1.1) arroja luz sobre la diversidad de enfoques que los gobiernos adoptan en relación con sus políticas económicas. Países líderes en este aspecto, como Hong Kong, Singapur y Suiza, destacan por crear entornos que fomentan el crecimiento y la inversión. Estos lugares se erigen como modelos de éxito, sugiriendo que una mayor autonomía económica puede ser un catalizador esencial para el desarrollo y la prosperidad.

En contraste, aquellos países con menores niveles de libertad económica podrían enfrentar obstáculos significativos en su camino hacia el desarrollo sostenible. La restricción de la libertad económica a menudo se traduce en limitaciones para la innovación, la inversión extranjera y el emprendimiento, elementos cruciales para construir economías robustas y sostenibles a largo plazo.

Este panorama diverso subraya la importancia de las políticas económicas en la determinación del éxito de una nación. La libertad económica no es simplemente un indicador abstracto, sino un componente vital que moldea el tejido económico y social de un país, influyendo en su capacidad para enfrentar desafíos y aprovechar oportunidades en el complejo escenario global.

4.2.2. Felicidad

La Intersección entre Libertad Económica y Bienestar Emocional

La conexión entre libertad económica y felicidad revela un fascinante panorama en la evaluación del progreso y el bienestar de una nación. La presencia de una correlación destacada entre altos índices de libertad económica y altos niveles de felicidad, como se evidencia en países ejemplares como Dinamarca y Suiza, sugiere una relación intrínseca entre el entorno económico y el bienestar emocional de la población.

La premisa fundamental aquí es que un entorno económico favorable, caracterizado por la libertad para emprender, innovar y participar en transacciones comerciales, no solo contribuye al crecimiento económico, sino que también influye directamente en la satisfacción y felicidad de las personas. Este fenómeno refleja una comprensión más holística del progreso de una sociedad, y destaca la importancia de considerar no solo los indicadores económicos tangibles, sino también los aspectos emocionales y subjetivos al evaluar el éxito de una nación.

En este contexto, la felicidad se convierte en un indicador de calidad de vida más allá de los parámetros tradicionales. Se subraya la idea de que una economía que permite a los individuos perseguir sus metas, disfrutar de oportunidades equitativas y participar en el proceso de desarrollo económico contribuye a un estado de satisfacción más elevado. Este enfoque más amplio refleja una comprensión avanzada del bienestar, reconociendo que la

prosperidad va más allá de las cifras financieras para abarcar la realización personal y la conexión emocional con la sociedad.

4.2.3. Esperanza de Vida

Un Vínculo Intrínseco con la Libertad Económica y la Felicidad

El análisis de la esperanza de vida revela un fascinante entrelazamiento entre la libertad económica, la felicidad y la longevidad de una población. Los países que destacan tanto en libertad económica como en felicidad, como Islandia y Japón, no solo experimentan una elevada satisfacción emocional, sino que también disfrutan de expectativas de vida notables.

Este patrón no es casualidad; más bien, sugiere una conexión profunda entre el entorno económico, el bienestar emocional y la salud general de una sociedad. La presencia de sistemas de salud robustos en estos países contribuye a una calidad de vida elevada y, por ende, a una mayor esperanza de vida. La inversión en infraestructuras sanitarias, acceso a servicios médicos de calidad y una atención médica preventiva eficiente se convierten en factores fundamentales que impulsan la longevidad de la población.

El mensaje clave aquí es que la libertad económica no solo impacta en la prosperidad financiera y la felicidad, sino que también tiene repercusiones positivas en la salud y la duración de la vida. Un entorno económico que fomente la inversión en sistemas de salud y promueva condiciones de vida saludables puede traducirse en una población más saludable y, por ende, en una mayor esperanza de vida. Este enfoque integral destaca la importancia de considerar la interconexión de diversos factores al abordar el desarrollo socioeconómico de una nación.

4.2.4. PBI per Cápita

La Riqueza Individual como Resultado de la Libertad Económica y la Felicidad Colectiva.

El análisis del Producto Bruto Interno (PBI) per cápita añade una dimensión esencial a la evaluación económica. Los países que ostentan niveles significativos de libertad económica y altos índices de felicidad suelen manifestar ingresos per cápita más elevados. Este fenómeno no es coincidencia, sino más bien una consecuencia directa de un entorno económico que fomenta el crecimiento y la inversión, junto con una población feliz y satisfecha.

La conexión entre la libertad económica y el PBI per cápita se revela como un ciclo virtuoso. La libertad económica impulsa la innovación, la inversión y el desarrollo empresarial, factores que, a su vez, contribuyen al aumento de la riqueza a nivel individual. A medida que los ciudadanos experimentan un crecimiento económico personal, esto alimenta un mayor sentido de bienestar y satisfacción.

Este fenómeno refuerza la idea de que la libertad económica va más allá de la mera acumulación de riqueza; es un catalizador que mejora la calidad de vida de los individuos. La prosperidad económica se traduce en mayores recursos disponibles para acceder a servicios esenciales, educación de calidad y oportunidades de desarrollo personal.

En resumen, el análisis del PBI per cápita confirma la interconexión entre libertad económica, felicidad y riqueza individual. La libertad económica no solo impulsa el crecimiento económico, sino que también contribuye directamente a la felicidad y al bienestar, generando un círculo virtuoso que eleva la calidad de vida de la población. Este enfoque holístico destaca la necesidad de políticas y prácticas que promuevan no solo el desarrollo económico, sino también la satisfacción y el florecimiento de las personas en sociedad.

4.2.5. Educación

La baja tasa de no escolarización y finalización de la educación en países caracterizados por altos niveles de libertad económica y felicidad subraya una conexión crucial entre el entorno económico, el bienestar emocional y la participación exitosa en el sistema educativo. Este fenómeno refleja cómo un entorno económico favorable no solo impacta la prosperidad material, sino que también influye positivamente en el acceso y la conclusión exitosa de la educación. La libertad económica crea las condiciones propicias para el desarrollo de sistemas educativos robustos y accesibles, donde los ciudadanos pueden disfrutar de la libertad para emprender, invertir y contribuir al crecimiento económico.

Países como Finlandia, Singapur y Canadá, que presentan altos índices de libertad económica y felicidad, también destacan por sus bajísimas tasas de no escolarización y altas tasas de finalización educativa. Estos ejemplos muestran cómo un entorno económico que fomenta la libertad puede traducirse en sistemas educativos exitosos y accesibles. En contraste, naciones con restricciones significativas en la libertad económica pueden enfrentar desafíos en el ámbito educativo. Por ejemplo, países con tasas más altas de no escolarización y finalización, como algunos en África subsahariana, podrían beneficiarse de políticas que promuevan la libertad económica para mejorar el acceso y la calidad de la educación.

La baja tasa de no escolarización y finalización también indica que la población, en general, está participando activamente en el sistema educativo. La educación, siendo un componente clave para el desarrollo individual y colectivo, se convierte en un pilar fundamental para el progreso sostenible de una sociedad. Noruega, Australia y Japón son ejemplos adicionales de naciones con altos índices de libertad económica, felicidad, bajas tasas de no escolarización y altas tasas de finalización. Estos países destacan por la integralidad de sus enfoques hacia el desarrollo socioeconómico, evidenciando cómo la libertad económica no solo se traduce en beneficios financieros, sino que también contribuye significativamente al acceso y la conclusión exitosa de una educación de calidad.

4.2.6. Desempleo

Una Perspectiva desde la Libertad Económica

La tasa de desempleo, como indicador clave de la salud económica, proporciona información valiosa cuando se examina en el contexto de la libertad económica. Países con altos niveles de libertad económica y bajos niveles de desempleo, como Hong Kong, destacan la efectividad de políticas económicas que promueven la libertad y la iniciativa empresarial.

La conexión entre la libertad económica y el bajo desempleo puede entenderse a través de varios mecanismos. En entornos donde las personas tienen la libertad de emprender y participar en actividades económicas sin restricciones excesivas, se fomenta la creación de empleo y el desarrollo de nuevas empresas. La flexibilidad del mercado laboral, impulsada por la libertad económica, permite adaptarse a las demandas cambiantes y facilita la generación de oportunidades laborales.

Además, la libertad económica también puede estar asociada a políticas que promueven la inversión y la innovación, creando un ambiente propicio para la expansión de industrias y la generación de empleo sostenible a largo plazo.

Matriz de Correlación: Relaciones Intrincadas entre Libertad Económica, Felicidad y Desarrollo Socioeconómico

La matriz de correlación agrega otra capa de comprensión al destacar la relación positiva entre libertad económica, felicidad y PBI per cápita. Este hallazgo refuerza la idea de que estos factores no son independientes, sino que están interconectados de manera intrínseca. La fortaleza de estas correlaciones sugiere que abordar estos aspectos de manera integral y coordinada es crucial para impulsar el progreso socioeconómico. Por ejemplo, políticas que fomenten la libertad económica no solo pueden tener un impacto directo en el crecimiento financiero, sino también en la satisfacción personal y el bienestar general de la población.

En resumen, la baja tasa de desempleo en países con alta libertad económica y los patrones revelados por la matriz de correlación resaltan la necesidad de estrategias integrales y coordinadas para abordar los desafíos socioeconómicos. El enfoque en la libertad económica no solo beneficia la salud del mercado laboral, sino que también contribuye a un panorama más amplio de desarrollo y bienestar.

4.2.7. Clustering y Calidad de Vida

Al aplicar algoritmos de clustering a los datos, se revelan patrones que agrupan países en categorías similares en términos de calidad de vida o bienestar general. Estos grupos no solo consideran indicadores económicos, sino también otros factores como la educación, las tasas de suicidio, el desempleo y la felicidad. La utilización de esta técnica proporciona una visión más completa y holística de las similitudes y disparidades entre las naciones.

Al explorar estos clusters, se puede identificar la presencia de patrones regionales o temáticos que van más allá de las métricas individuales. Este enfoque ayuda a comprender mejor las dinámicas complejas que influyen en la calidad de vida, permitiendo así un abordaje más preciso y específico en las estrategias de desarrollo y mejora de la sociedad.

En resumen, el clustering proporciona una herramienta valiosa para discernir tendencias y similitudes entre países, permitiendo a los responsables de la toma de decisiones diseñar intervenciones más efectivas y adaptadas a las necesidades específicas de cada grupo.

5. Discusión

5.1. Comparación con Objetivos:

Al evaluar en qué medida se cumplieron los objetivos planteados en el proyecto, se puede observar que se lograron con éxito varios hitos fundamentales:

- **Identificación de Patrones:** El objetivo de identificar patrones y agrupamientos significativos en datos socioeconómicos se cumplió mediante el uso de técnicas de clustering. Los clusters revelaron conexiones entre diferentes indicadores, proporcionando una comprensión más profunda de las dinámicas globales.
- **Visualización Efectiva:** La creación de mapas interactivos y gráficos facilitó la interpretación de los datos. Las visualizaciones ayudaron a comunicar eficazmente las variaciones en la libertad económica, felicidad, esperanza de vida, PIB per cápita, tasas de suicidio y datos educativos a nivel mundial.
- **Relaciones Intrínsecas:** La matriz de correlación examinó las relaciones estadísticas entre variables clave, reforzando la idea de interconexión entre libertad económica, felicidad y desarrollo socioeconómico.

5.2. Desafíos y Limitaciones:

A pesar del éxito en alcanzar los objetivos, se enfrentaron algunos desafíos y limitaciones durante el desarrollo del proyecto:

- **Calidad de los Datos:** La calidad de los resultados dependió directamente de la calidad de los datos disponibles. Limitaciones en la precisión o integridad de los conjuntos de datos podrían haber afectado la robustez de las conclusiones.
- **Selección de Variables:** La elección de las variables a analizar y su relevancia podría influir en los resultados. Se reconoce que, aunque se abordaron múltiples indicadores, aún existen otros factores importantes que podrían haber sido considerados para una comprensión más completa.
- **Complejidad de las Relaciones Causales:** Aunque se identificaron correlaciones, establecer relaciones causales directas entre la libertad económica, la felicidad y otros indicadores puede ser complejo. Factores externos y variables no consideradas pueden influir en las relaciones.

6. Conclusiones

6.1. Reflexiones Finales:

El análisis integral de datos socioeconómicos utilizando Machine Learning y Big Data ha revelado conexiones profundas entre la libertad económica, felicidad, salud y educación. La comprensión de estas interrelaciones proporciona una base sólida para la toma de decisiones informadas en el ámbito gubernamental y social.

6.2. Aplicación de Conocimientos:

La aplicación de técnicas de Machine Learning, como el clustering, y el procesamiento de Big Data permitieron una exploración detallada de patrones y relaciones en datos complejos. La visualización efectiva y la interpretación de la matriz de correlación demostraron cómo estos conceptos avanzados pueden mejorar significativamente la comprensión de fenómenos socioeconómicos a escala global.

6.3. Sugerencias para Futuras Investigaciones:

Para futuras investigaciones, se sugiere explorar más a fondo ciertos aspectos, como:

- **Inclusión de más Variables:** La adición de más variables podría enriquecer aún más el análisis, considerando factores culturales, medio ambiente y desarrollo tecnológico.
- **Temporalidad de los Datos:** La introducción de datos temporales permitiría analizar tendencias a lo largo del tiempo y comprender mejor la dinámica de cambio en los indicadores.
- **Estudios de Caso Detallados:** Realizar estudios de caso detallados en países específicos para comprender mejor los factores contextuales que pueden influir en las relaciones identificadas en este proyecto.
- **Validación con Expertos:** Validar los resultados con expertos en economía, salud y educación para asegurar la precisión y relevancia de las conclusiones.

7. Entrega de Código y Documentación

7.1. Enlace al Código en GitHub:

[Proyecto_final_bigdata/Proyecto_final_MLBG.ipynb at main · Matiitwtcp/Proyecto_final_bigdata \(github.com\)](https://github.com/Matiitwtcp/Proyecto_final_bigdata)

7.2. Enlace a los Datasets en Drive:

<https://drive.google.com/drive/folders/1Ep7ujDM4PuruSwECx-9Ft4xpWpyfSleF?usp=sharing>

8. Referencias

8.1. Fuentes de datos, bibliografía y recursos utilizados.

- Datos económicos: Instituto Fraser.
- Datos educativos: UNICEF, UNESCO.
- Datos de salud: Organización Mundial de la Salud.
- Herramientas y bibliotecas: Documentación oficial de pandas, geopandas, plotly, etc.

- anexo 1:

- Libertad Económica (2016):
<https://www.kaggle.com/datasets/gsutters/economic-freedom/data>
Utiliza cinco áreas: tamaño del gobierno, estructura legal y seguridad de los derechos de propiedad, acceso a dinero sólido, libertad para comerciar internacionalmente y regulación del crédito, del trabajo y los negocios.
- Felicidad global (2021):
<https://www.kaggle.com/code/noobiedatascientist/explaining-happiness>
Utiliza diferentes conjuntos de datos como: PIB per cápita, crecimiento demográfico, esperanza de vida, edad media, consumo de carne, proporción de sexos, tasa de suicidio, urbanización, tasa de fertilidad.
- Esperanza de vida, PBI y tasa de suicidios (2020):
<https://www.kaggle.com/datasets/daniboy370/world-data-by-country-2020>
Obtiene y organiza información de
https://en.wikipedia.org/wiki/Category:Lists_of_countries
- Tasa de NO Escolaridad y de Finalización:
<https://www.kaggle.com/datasets/nelgiriyeewithana/world-educational-data>
Tasa de finalización - Base de datos mundial UNICEF (abril de 2021):
Tasa de niños y jóvenes sin escolarizar - Instituto de Estadística de la UNESCO (marzo de 2021).

- anexo 2:

<https://colab.google/>, <https://pandas.pydata.org/>, <https://geopandas.org/>, <https://numpy.org/>,
<https://matplotlib.org/>, <https://plotly.com/python/>, <https://scikit-learn.org/>.