

LA FELICIDAD EN EL MUNDO

Guillem Pañella Valiente

Tutores: Alberto Romero, Ivan Cordero, Lucas Zamora
Data Science Online Madrid Mayo 2024 - Diciembre 2024

ÍNDICE DE CONTENIDO

1. ANÁLISIS EXPLORATORIO DE DATOS (EDA):

- 1.1. Explicación motivación base de datos y diferentes variables
- 1.2. Hipótesis
- 1.3. Acondicionamiento bases de datos y tratamiento de nulos
- 1.4. Añadir variables necesarias para análisis
- 1.5. Análisis univariante
 - 1.5.1. Distribución de cada variable
 - 1.5.2. Boxplot de cada variable
 - 1.5.3. Estadísticas descriptivas de cada variable
 - 1.5.4. Evolución de países más felices y menos felices
 - 1.5.5. Análisis tablas extraídas
- 1.6. Análisis bivalente
 - 1.6.1. Análisis por continente
 - 1.6.2. Matriz correlación por país
 - 1.6.3. Análisis de las matrices de correlación

2. CONCLUSIONES

1. ANÁLISIS EXPLORATORIO DE DATOS (EDA)

1.1. Explicación motivación base de datos y diferentes variables

En cualquier departamento de Recursos Humanos en el que quieran saber si sus trabajadores son felices o no, necesitan saber qué es lo que más feliz hace a su sociedad, a la población de su país para convertirse en una empresa atractiva para trabajar en ella. Buscar ser una empresa en la que dé gusto trabajar y que las personas quieran trabajar en ella tiene que conocer cómo es la felicidad en el mundo, porque las personas son felices y como este indicador objetivo de felicidad varía según el país al que tengamos la sede.

Sin entrar aún en detalles ni diferencias entre países, pero lo que es obvio es que no será feliz por las mismas cosas una persona que vive en Canadá que en España que en Noruega o que en Afganistán. Las diferencias socioeconómicas y políticas van a determinar porque una persona dice ser feliz en un lugar o no. De la misma manera, lo que hace infeliz a las personas también variará según el país del que provenga una persona.

Por último, las diferencias también se pueden hacer notorias según el momento en que se pregunte si eres feliz o no. Sin ir más lejos, una pandemia a nivel mundial como el COVID afectó en la felicidad de muchísimas personas y no de la misma manera.

Por ello, en este Análisis Exploratorio de Datos (EDA) se utilizará la base de datos de World Happiness Report. El World Happiness Report se realiza con una asociación de Gallup, el Centro de Investigación del Bienestar de Oxford, la Red de Soluciones para el Desarrollo Sostenible de las Naciones Unidas y el Consejo Editorial de la WHR. El informe se produce bajo el control editorial del Consejo Editorial de WHR.

Por otro lado, “el World Happiness Report refleja una demanda mundial de más atención a la felicidad y el bienestar como criterios para las políticas gubernamentales. Revisa el estado de la felicidad en el mundo actual y muestra cómo la ciencia de la felicidad explica las variaciones personales y nacionales en la felicidad” (WHR en su página web, 2024).

También parece pertinente explicar muy bien las variables que vamos a utilizar durante el EDA y que se entiendan bien cuales son sus índices y medidas.

1. Life Ladder

Descripción: El "Life Ladder" (Escalera de la Vida) es una medida de bienestar subjetivo.

Escala: 0 a 10.

Interpretación:

0: Representa la peor vida posible según la percepción del individuo.

10: Representa la mejor vida posible según la percepción del individuo.

Cálculo: Se basa en la pregunta de la encuesta Gallup World Poll donde se pide a los encuestados que imaginen una escalera con peldaños numerados del 0 al 10 y que indiquen en qué peldaño se sienten actualmente.

2. GDP per Capita

Descripción: Producto Interno Bruto (PIB) per cápita ajustado por paridad de poder adquisitivo (PPA).

Medición: Se mide en dólares internacionales ajustados por PPA y se trata como un logaritmo, durante el EDA se genera una variable para que se vea el PIB per cápita real

Cálculo: Utiliza datos del Banco Mundial y otras fuentes económicas para calcular el PIB per cápita, ajustando las diferencias en el costo de vida entre países.

3. Social Support

Descripción: Medida de la percepción de apoyo social.

Interpretación: Escala del 0 al 1

Cálculo: Basada en la pregunta de la encuesta Gallup World Poll: "Si te encuentras en dificultades, ¿tienes familiares o amigos con quienes puedas contar para recibir ayuda cuando la necesites, o no?". La proporción de personas que responden "sí" se utiliza para esta variable.

4. Healthy Life Expectancy

Descripción: Esperanza de vida saludable al nacer.

Medición: Años.

Cálculo: Combina datos de la Organización Mundial de la Salud (OMS) y otras fuentes de salud pública para calcular los años de vida esperados en buena salud.

5. Freedom to Make Life Choices

Descripción: Medida de la percepción de libertad para tomar decisiones sobre la vida.

Interpretación: Escala del 0 al 1

Cálculo: Basada en la pregunta de la encuesta Gallup World Poll: "¿Estás satisfecho o insatisfecho con tu libertad para elegir lo que haces con tu vida?". La proporción de personas que responden "satisfecho" se utiliza para esta variable.

6. Generosity

Descripción: Medida de la generosidad de la población.

Interpretación: Escala del 1 al -1

Cálculo: Basada en la pregunta de la encuesta Gallup World Poll: "¿Has donado dinero a una organización benéfica en el último mes?". La proporción de personas que responden "sí" se ajusta por el PIB per cápita.

7. Perceptions of Corruption

Descripción: Medida de la percepción de corrupción en el gobierno y negocios.

Interpretación: Escala del 0 al 1 (como más cerca de 1 más corrupción percibida)

Cálculo: Basada en la pregunta de la encuesta Gallup World Poll: "¿Está extendida la corrupción en el gobierno o no?" y "¿Está extendida la corrupción en los negocios o no?". La proporción de personas que responden "sí" a cualquiera de las preguntas se utiliza para esta variable.

8. Positive Affect

Descripción: Medida de las emociones positivas.

Interpretación: Escala del 0 al 1

Cálculo: Basada en preguntas de la encuesta Gallup World Poll sobre si los encuestados experimentaron ciertas emociones positivas (risa, disfrute, etc.) el día anterior a la encuesta. La proporción de respuestas positivas se utiliza para esta variable.

9. Negative Affect

Descripción: Medida de las emociones negativas.

Interpretación: Escala del 0 al 1

Cálculo: Basada en preguntas de la encuesta Gallup World Poll sobre si los encuestados experimentaron ciertas emociones negativas (preocupación, tristeza, etc.) el día anterior a la encuesta. La proporción de respuestas positivas se utiliza para esta variable.

1.2. Hipótesis

Una vez planteada la utilización de la base de datos y el por qué de su utilización, pasamos al planteamiento de las diferentes hipótesis del trabajo:

1. Qué variables tienen un MAYOR efecto sobre el índice de felicidad o LIFE LADDER
2. Comparar diferentes países y la variable que tienen más efecto en el índice de felicidad o LIFE LADDER

3. Las variables de generosidad y apoyo social han crecido desde la pandemia en el año 2020.
4. Los países que lideran una variable no tienen porqué ser los más felices
5. Los 5 países MÁS felices del mundo tienen una ideología política basada en la socialdemocracia (busca apoyar las intervenciones estatales, tanto económicas como sociales, para promover mayor equidad económica e igualdad social en el marco de una economía capitalista)
6. Los países mucho menos desarrollados económicos (PIB más bajo) son los que tienen menor LIFE LADDER

1.3. Acondicionamiento bases de datos y tratamiento de nulos

1.4. Añadir variables necesarias para análisis

En estos dos puntos son más de acondicionamiento de nuestra base de datos sobre el entorno en el que trabajaremos qué es python desde un notebook para ir resolviendo todas las hipótesis.

La base de datos encontrada estaba bastante limpia en cuanto al contenido y se tuvieron que hacer solo tres tareas importantes para poder trabajar en un futuro:

- a) Utilizar la media para sustituir los nulos. Al ser todo variables numéricas nos servía para no perder ningún valor dentro de nuestra base de datos.
- b) Sustituir las comas por puntos ya que muchas variables si utilizamos el método info() nos informaba que la mayoría de variables eran objetos. Es por ello que planteamos una función que lo sustituye con el método replace.
- c) Creación de variables necesarias para el análisis:
 - i) Variable continente para saber a qué continente pertenece cada país
 - ii) Variable PIB per cápita para saber y poder clasificar los países según sus ingresos con la clasificación planteada por el Banco Mundial
 - iii) Variable categoría de ingresos a partir de la mencionada clasificación cada país tenía un PIB per cápita real (no un logaritmo) y su categoría de ingreso
 - iv) Variable temporal pre covid (antes del 2020) y post-covid (después del 2020)

La clasificación a la que hacemos referencia para categorizar los ingresos por país es la siguiente:

- Ingresos Altos: PIB per cápita superior a \$13,205 USD.
- Ingresos Medios:
 - Medios Altos: PIB per cápita entre \$4,256 y \$13,205 USD.
 - Medios Bajos: PIB per cápita entre \$1,046 y \$4,255 USD.

- Ingresos Bajos: PIB per cápita inferior a \$1,046 USD.

1.5. Análisis univariante

1.5.1. Distribución de cada variable

1.5.2. Boxplot de cada variable

1.5.3. Estadísticas descriptivas de cada variable

En este apartado se hace uso de diferentes funciones con python para poder realizar el análisis inicial univariante. Estas funciones nos permiten entender a la perfección la distribución de nuestras variables con un histograma y entender bien los extremos que tiene cada una de las variables con un boxplot por cada una.

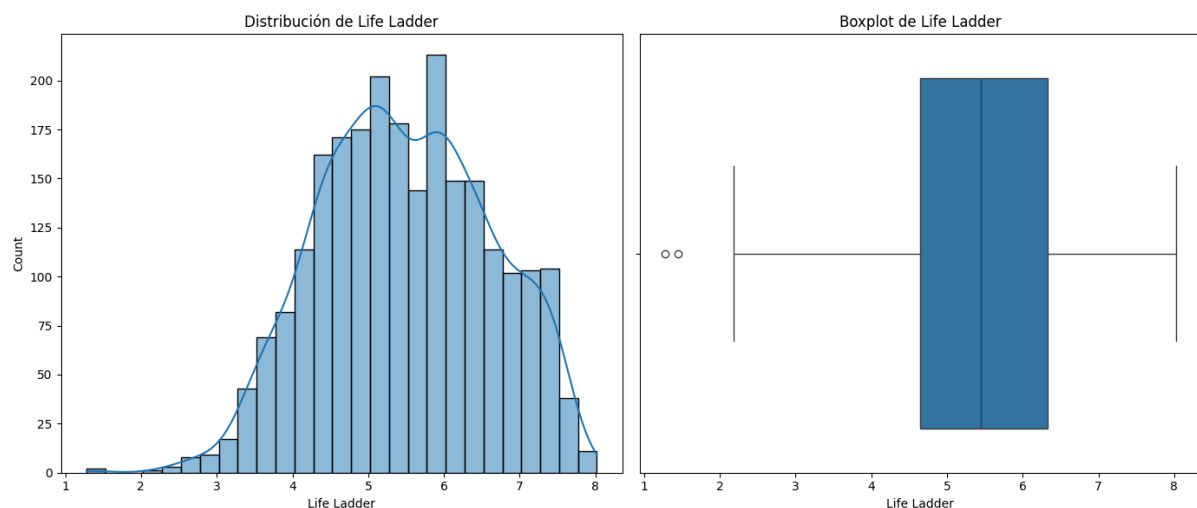
Además dentro de la función también se incluye que pueda devolver los principales estadísticos de esa variable incluidos en el método describe() como la media, la desviación estándar, la mediana y los cuartiles y los valores máximos y mínimos de esa variable.

La función devuelve una figura por cada variable donde podemos ver los gráficos mencionados. Se muestra en la memoria el análisis y lo que devuelve la función planteada de la variable life ladder o índice de felicidad que es la variable más importante de nuestro análisis.

Análisis univariante para Life Ladder:

```
count    2363.000000
mean      5.483566
std       1.125522
min       1.281000
25%       4.647000
50%       5.449000
75%       6.323500
max       8.019000
```

Figura 1: histograma y boxplot de la variable Life Ladder



Con esta información de cada variable nos da una perspectiva mucho más completa de cómo es la muestra de cada una de las variables y como se ha respondido durante los últimos 17 años la encuesta sobre la World Happiness.

1.5.4. Evolución de países más felices y menos felices

En este caso, se genera una función para que se puedan ver claramente los países más felices y menos de los últimos 6 años (2017 hasta 2023). Sigue siendo dentro del apartado del análisis univariante ya que estamos ante los máximos y mínimos de cada uno de los años analizados. Nos interesa ver que tipología de países están en cada top/flop y ver qué tienen en común para que luego cuando se crucen más de una variable se pueda entender mejor.

Para ver algunas ideas ya de los datos que nos arroja el World Happiness Report, vemos la tabla de 2017, 2019, 2021, 2023:

Figura 2: Tablas de los top5 países más felices y los flop5 países menos felices (años 2017, 2019, 2021, 2023)

Resultados para el año 2017:					
	Country name	Year	Max Life Ladder	Min Life Ladder	
40	Finland	2017	7.788	7.788	
33	Denmark	2017	7.594	7.594	
100	Norway	2017	7.579	7.579	
54	Iceland	2017	7.476	7.476	
125	Switzerland	2017	7.474	7.474	
..	
128	Tanzania	2017	3.347	3.347	
144	Yemen	2017	3.254	3.254	
110	Rwanda	2017	3.108	3.108	
120	South Sudan	2017	2.817	2.817	
0	Afghanistan	2017	2.662	2.662	

Resultados para el año 2019:

	Country name	Year	Max Life Ladder	Min Life Ladder
39	Finland	2019	7.780	7.780
122	Switzerland	2019	7.694	7.694
31	Denmark	2019	7.693	7.693
52	Iceland	2019	7.533	7.533
98	Norway	2019	7.442	7.442
..
141	Zambia	2019	3.307	3.307
108	Rwanda	2019	3.268	3.268
53	India	2019	3.249	3.249
142	Zimbabwe	2019	2.694	2.694
0	Afghanistan	2019	2.375	2.375

Resultados para el año 2021:

	Country name	Year	Max Life Ladder	Min Life Ladder
32	Finland	2021	7.794	7.794
26	Denmark	2021	7.699	7.699
49	Israel	2021	7.578	7.578
43	Iceland	2021	7.565	7.565
102	Sweden	2021	7.439	7.439
..
44	India	2021	3.558	3.558
121	Zimbabwe	2021	3.155	3.155
120	Zambia	2021	3.082	3.082
0	Afghanistan	2021	2.436	2.436
61	Lebanon	2021	2.179	2.179

Resultados para el año 2023:

	Country name	Year	Max Life Ladder	Min Life Ladder
38	Finland	2023	7.699	7.699
50	Iceland	2023	7.562	7.562
31	Denmark	2023	7.504	7.504
27	Costa Rica	2023	7.384	7.384
89	Netherlands	2023	7.255	7.255
..
108	Sierra Leone	2023	3.467	3.467
26	Congo (Kinshasa)	2023	3.383	3.383
13	Botswana	2023	3.332	3.332
74	Malawi	2023	3.272	3.272
0	Afghanistan	2023	1.446	1.446

Aun sin entrar en un análisis profundo, podemos observar que Finlandia se repite en los 4 años visualizados como el país más feliz y países con vivencias de guerra como Afghanistan y Líbano son los que se declaran como menos felices dentro del WHR.

Sin salir todavía del análisis univariante, se ha buscado saber también cuáles son los países más ricos de forma absoluta y ver si realmente estos también aparecían en los top5 más felices. Es decir, buscar si “el dinero da la felicidad” se cumple en las respuestas del WHR.

Con los métodos de nsmallest y nlargest conseguimos la siguiente tabla:

Figura 3: Tabla de países más ricos y más pobres del WHR en términos globales

10 países con mayor PIB per capita:				
	Country name	year	PIB per capita	Life Ladder
970	Ireland	2023	117712.442342	6.817
1244	Luxembourg	2016	116308.334530	6.967
1248	Luxembourg	2022	115497.019105	7.228
1245	Luxembourg	2017	115035.953774	7.061
1247	Luxembourg	2019	114576.729021	7.404
1249	Luxembourg	2023	114576.729021	7.016
1238	Luxembourg	2010	114347.804564	7.097
1246	Luxembourg	2018	114119.337498	7.243
969	Ireland	2022	113891.326910	6.870
1242	Luxembourg	2014	113323.291547	6.891
10 países con menor PIB per capita:				
	Country name	year	PIB per capita	Life Ladder
2291	Venezuela	2019	251.388613	5.081
2290	Venezuela	2018	378.039996	5.006
2289	Venezuela	2017	381.076446	5.071
312	Burundi	2018	740.258907	3.775
309	Burundi	2009	801.912901	3.792
310	Burundi	2011	807.545984	3.706
463	Congo (Kinshasa)	2009	811.593825	3.984
308	Burundi	2008	812.405825	3.563
370	Central African Republic	2016	818.112616	2.693
311	Burundi	2014	831.307697	2.905

Como podemos observar en la anterior tabla, en la parte de 10 países con mayor PIB solo aparecen Irlanda y Luxemburgo. En estos dos casos no los hemos visto aparecer en ningún top de felicidad. Por lo tanto, sin aún cruzar las variables para ver qué impacto tiene el PIB per cápita en el índice de felicidad nos dice que ser el país MÁS rico no tiene porque hacerte el país MÁS feliz sino que el tener variables más altas ayuda a ser de los países más felices cómo era el caso de Finlandia.

En el caso de los países más pobres en PIB per cápita pasa algo parecido que con los más ricos. Afecta en que el life ladder no sea superior a 5, sin embargo, no los convierte en los países menos felices. Añadir también que, en términos generales, en los últimos años los países se han enriquecido. Esto se puede observar en que el top10 de países ricos la gran mayoría de países más ricos son en los últimos 10 años. Sin embargo, los registros de PIB per cápita más bajos son en la gran mayoría antes del 2017, es decir, hace 5 años ya.

1.5.5. Análisis tablas extraídas

De forma global y para cerrar el análisis univariante vamos a detectar las diferentes ideas que han salido de hacer los diferentes gráficos y tablas en este apartado.

En primer lugar, las respuestas dentro del índice de felicidad o life ladder se concentran en gran medida entre el 5 y el 7. Situarse por encima o por debajo de estos valores ya te sitúa dentro de los más felices y menos felices dentro de lo que es el ranking anual de cada uno de los años analizados.

Con el PIB per cápita pasa algo distinto ya que la forma que coge su histograma es de una forma conocida como exponencial. De hecho, si has visto cómo se distribuye la riqueza en un país como Estados Unidos habrás visto una distribución parecida. Se trata de que unos pocos tienen un porcentaje de riqueza muy alto. El conocido 1% más rico de Estados Unidos tiene el 40% de la riqueza total del país. Hay, entonces, una gran desigual repartición de la riqueza.

De esta forma, en el caso de la distribución de esta variable vemos que hay muy pocos países que son muy ricos, pero los que lo son, lo son mucho. Es el caso, como hemos observado, de Luxemburgo, Irlanda, países nórdicos que superan los 100.000 dólares cuando la categoría de ingresos altos del Banco Mundial marca los 13.205\$. Además, comprobamos estos valores extremos también con el boxplot de esa misma variable.

Sin entrar aún en el detalle de lo que pueden hacer las otras variables para hacer más felices a los países sí que el análisis univariante clarifica que el resto de variables se distribuyen de una similar forma que la variable de Life Ladder o índice de felicidad.

1.6. Análisis bivalente

Una vez se han entendido bien las diferentes variables de forma individual, como se distribuyen y cómo han evolucionado en los diferentes años llega el momento de cruzar diferentes variables para entender bien cómo afectan las diferentes variables a ser feliz. Es decir, a partir de aquí se entenderá si el PIB per cápita afecta a ser feliz o es la esperanza de vida o la sensación de corrupción. También si existen diferencias según países o continentes en este efecto de las variables en ser felices.

1.6.1. Análisis por continente

Antes de ver la matriz de correlación de las variables por continente, es pertinente para hacernos una idea de como se relacionan todas las variables con el life ladder un gráfico de dispersión o más conocido en inglés como scatter plot.

Figura 4: scatter plot de la variable life ladder con la de log PIB per cápita

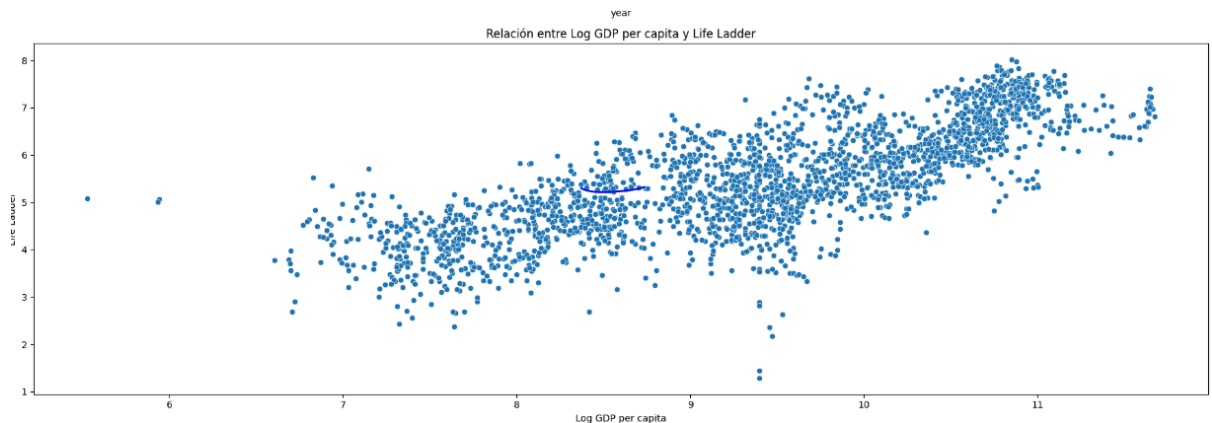


Figura 5: scatter plot de la variable life ladder con la libertad de tomar decisiones

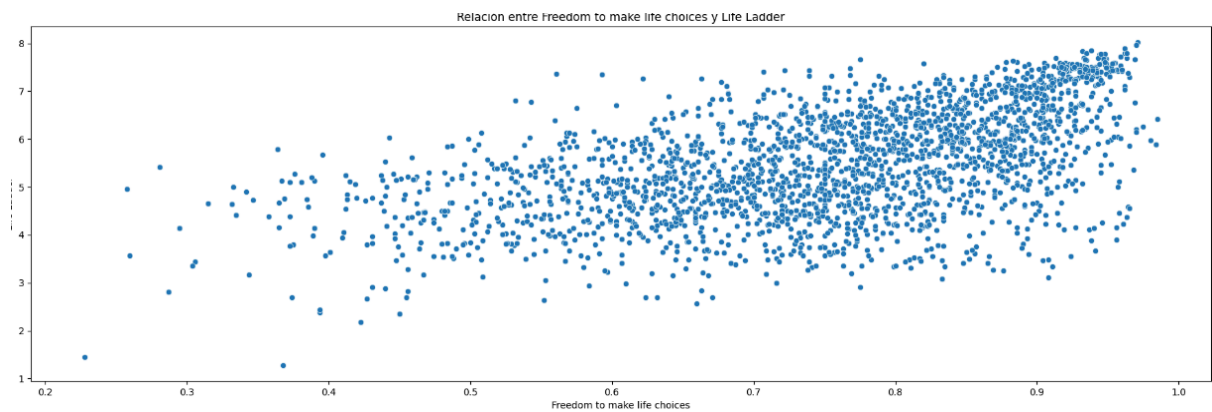
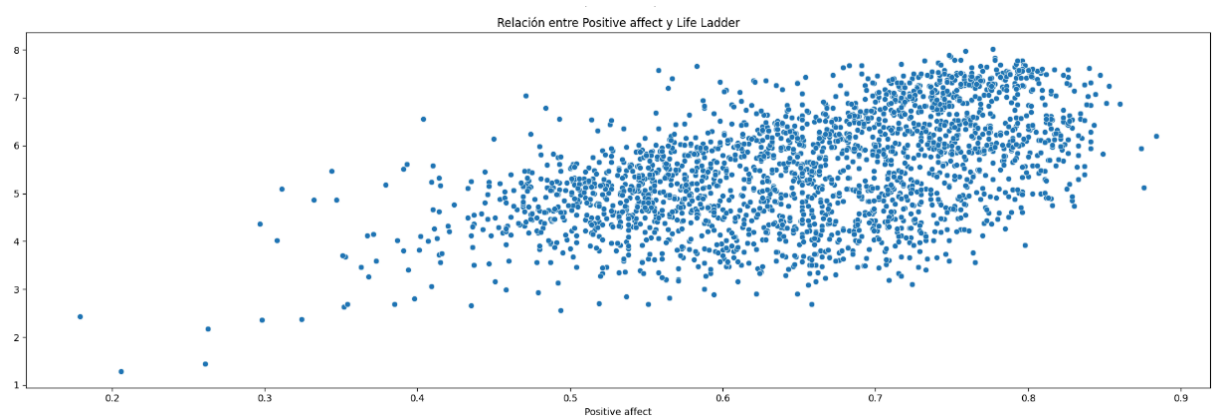


Figura 6: scatter plot de la variable life ladder con la afectos positivos



En estos gráficos ya se pueden ver tendencias y relaciones entre las variables de la WHR con nuestra variable Life Ladder que mide la felicidad. El ejemplo más claro y visible es el de la figura 4 en el que la recta positiva de relación entre ambas variables y sus puntos se puede hacer más claramente. En el momento en que se tiene más dinero, más sube el índice de felicidad o life ladder aunque nos referiremos más tarde a esto.

Figura 7: Tabla de correlación en el continente de Asia

	Variable	Correlation	continente
0	Life Ladder	1.000000	Asia
1	Log GDP per capita	0.679979	Asia
2	Social support	0.658545	Asia
3	Healthy life expectancy at birth	0.559910	Asia
4	Freedom to make life choices	0.394555	Asia
5	Generosity	0.062017	Asia
6	Perceptions of corruption	-0.220007	Asia
7	Positive affect	0.505172	Asia
8	Negative affect	-0.390224	Asia

Figura 8: Tabla de correlación en el continente de Europa

	Variable	Correlation	continente
0	Life Ladder	1.000000	Europe
1	Log GDP per capita	0.821318	Europe
2	Social support	0.658877	Europe
3	Healthy life expectancy at birth	0.698284	Europe
4	Freedom to make life choices	0.776565	Europe
5	Generosity	0.537987	Europe
6	Perceptions of corruption	-0.789912	Europe
7	Positive affect	0.838792	Europe
8	Negative affect	-0.547754	Europe

Figura 9: Tabla de correlación en el continente de Oceania

	Variable	Correlation	continente
0	Life Ladder	1.000000	Oceania
1	Log GDP per capita	-0.444436	Oceania
2	Social support	0.500448	Oceania
3	Healthy life expectancy at birth	-0.574027	Oceania
4	Freedom to make life choices	0.541432	Oceania
5	Generosity	0.364777	Oceania
6	Perceptions of corruption	-0.263180	Oceania
7	Positive affect	0.571961	Oceania
8	Negative affect	-0.245202	Oceania

En este análisis vemos cómo varía la relación que hay entre todas las variables y nuestra variable independiente, el índice de felicidad o life ladder. En esta memoria ponemos la de

tres continentes pero para que se haga más extenso, en el código se pueden consultar visualmente los 6 continentes a partir de la función que se plantea.

Entrando en el análisis en todos los continentes menos Oceanía el PIB per cápita es la variable que mayor fuerza demuestra en la tabla de correlación. Además se trata de una correlación en todos los casos positiva, es decir, cuando sube una, la otra también sube.

Por otro lado, es interesante ver el caso del continente de Oceanía que a parte de que la correlación entre PIB per cápita y life ladder sea negativa no es tan fuerte como la de esperanza de vida o las emociones positivas. Es decir, en Oceanía eres más feliz cuanto menos PIB tiene el país, cuanto menor es la esperanza de vida mayor es el índice de felicidad y, por último, las emociones positivas afectan en que el índice de felicidad crezca. Es con diferencia el continente con variables más diferentes al resto y que más sorprenden. Adicionalmente, es interesante como Europa es el continente que tiene una correlación más fuerte (0,82) entre las variables de Life Ladder y PIB per cápita superando al segundo continente que mayor fuerza tiene esa correlación que es North America (0,73).

El resto de variables tiene una correlación con menos fuerza pero una que parece destacable es la del apoyo social en Asia que tiene más fuerza que en el resto de continentes y luego veremos con más detalle.

Siguiendo analizando el cruce de diferentes variables una de las hipótesis de este trabajo era también analizar qué había pasado en el momento de post pandemia con las variables de Generosidad y de apoyo social. Para ello, con la creación de una función de gráfico de líneas (line plot) nos permitirá evidenciar estas diferencias por continente.

Figura 10: gráfico de líneas sobre la evolución por continente de la variable social support (2005-2023)

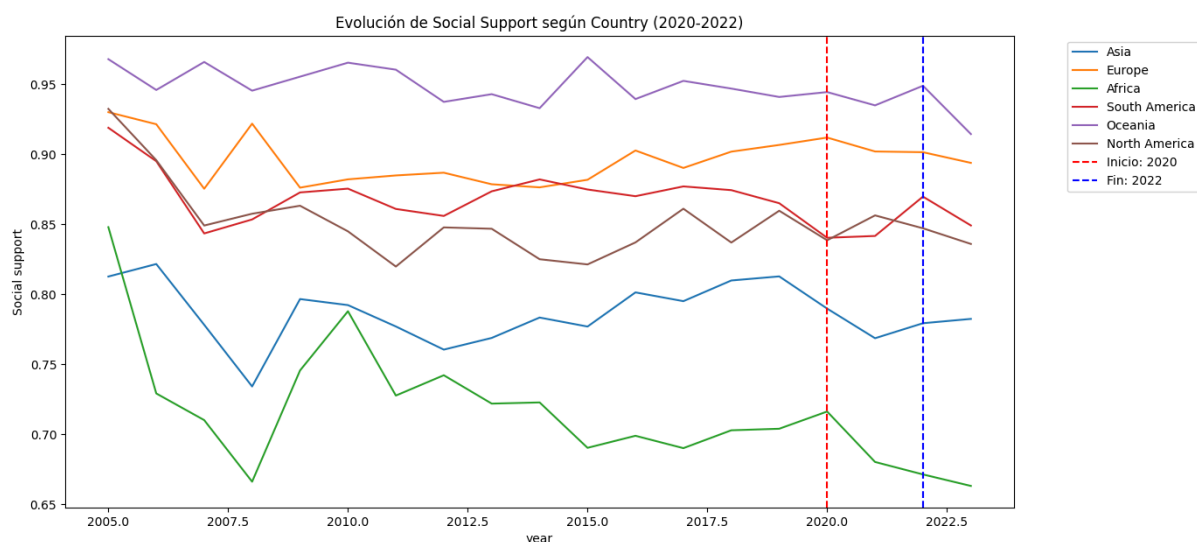
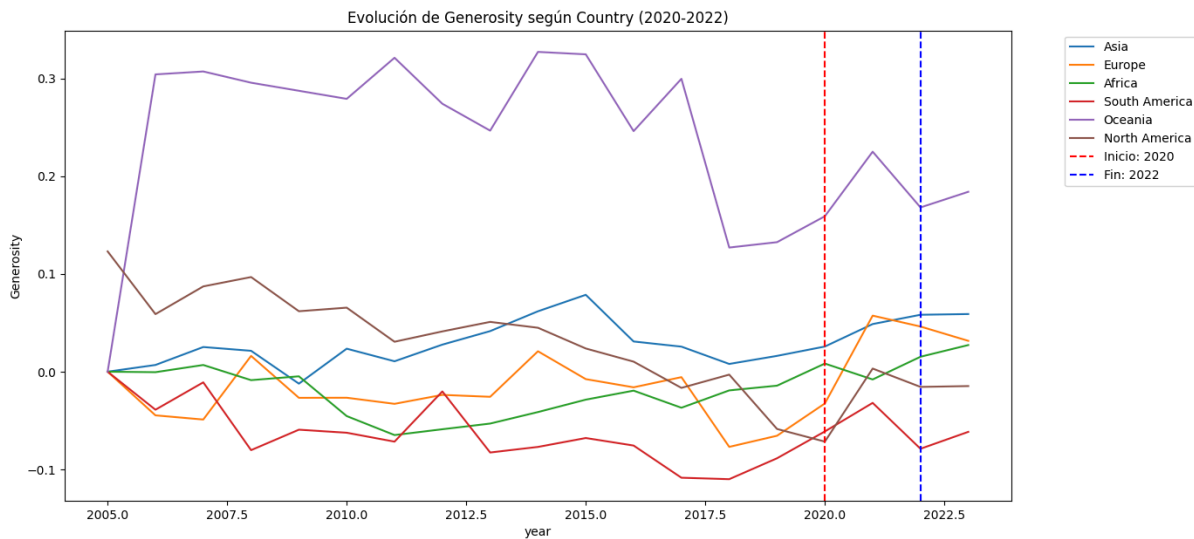


Figura 11: gráfico de líneas sobre la evolución por continente de la variable generosidad (2005-2023)



La relación de estas variables con la evolución de los años era importante para entender si las personas con la llegada de una pandemia que cambiaba mucho la forma de relacionarnos durante 2 años podían afectar. Hay una variable que claramente ve cómo crece su promedio y es la generosidad. Si nos fijamos en la figura 11 todos los continentes tienen una subida de la variable generosidad en ese periodo de 2020 a 2022. No obstante, menos en Asia, el resto de continentes tienen una subida en 2021 y ven como en 2022 vuelve a bajar, creando una figura así de montaña de subida y bajada muy clara visualmente. En Asia la subida es menos notoria pero más progresiva y sin parar.

Por otro lado, la variable de apoyo social se comporta diferente en los diferentes continentes. En África, esta variable baja cada año a partir de 2020 y no cambia la tendencia igual que en Europa aunque la bajada es menos notoria en Europa. En Asia y Oceanía pasa lo inverso a lo que pasaba con la generosidad, es decir, baja en 2021 la variable de apoyo social pero luego vuelve a subir en 2022 para cambiar la tendencia hacia arriba. Por último, la South y North America crecen en esa variable de apoyo social durante los años post covid y en la figura 11 se denota esa tendencia.

1.6.2. Matriz correlación por país y modelo de regresión lineal

En este apartado se ha hecho algo parecido a lo del apartado anterior pero entrando a un nivel de granularidad y detalle más alto. De esta forma, con una función de creación de matriz de correlación pero por países podremos ver diferentes países que interesen para el análisis de este trabajo y ver las diferencias. Posteriormente en el siguiente apartado

entraremos al detalle de estas correlaciones y entender muy bien esas diferencias y la fuerza de esas correlaciones.

En este caso la función nos da la correlación de todas las variables con la variable independiente de índice de felicidad o life ladder pero la da de forma global de todos los años analizados (2005 a 2023) y no distingue en ese sentido. No obstante, con estas tablas se detectarán ya bastantes diferencias teniendo como referencia cada uno de los continentes analizados ya en el apartado anterior.

A modo aclaratorio, se lanzará la función para los siguientes países: España, Luxemburgo, Noruega, Irlanda, Italia, Portugal, China, Japón, India, Tailandia, Estados Unidos, Venezuela, Argentina, Brasil, Australia. Como se puede observar, son representación de cada uno de los continentes y con diferencias que ya se ha podido ir detectando durante el análisis realizado y que nos pueden interesar. En la memoria no pondremos las figuras de todos los países pero en el archivo del notebook se podrá acceder a verlas todas.

Figura 12: Tabla de correlación del país España con variable independiente Life Ladder y con el resto de variables del análisis

	Variable	Correlation	Country
0	Life Ladder	1.000000	Spain
1	Log GDP per capita	0.277633	Spain
2	Social support	0.455667	Spain
3	Healthy life expectancy at birth	-0.457430	Spain
4	Freedom to make life choices	0.673874	Spain
5	Generosity	-0.039403	Spain
6	Perceptions of corruption	-0.405305	Spain
7	Positive affect	0.456601	Spain
8	Negative affect	-0.728845	Spain

Figura 13: Tabla de correlación del país Luxemburgo con variable independiente Life Ladder y con el resto de variables del análisis

	Variable	Correlation	Country
0	Life Ladder	1.000000	Luxembourg
1	Log GDP per capita	0.337926	Luxembourg
2	Social support	-0.194862	Luxembourg
3	Healthy life expectancy at birth	0.174775	Luxembourg
4	Freedom to make life choices	-0.193125	Luxembourg
5	Generosity	-0.550197	Luxembourg
6	Perceptions of corruption	-0.062071	Luxembourg
7	Positive affect	-0.130529	Luxembourg
8	Negative affect	0.227229	Luxembourg

Figura 14: Tabla de correlación del país Norway con variable independiente Life Ladder y con el resto de variables del análisis

	Variable	Correlation	Country
0	Life Ladder	1.000000	Norway
1	Log GDP per capita	-0.685863	Norway
2	Social support	0.018915	Norway
3	Healthy life expectancy at birth	-0.408240	Norway
4	Freedom to make life choices	0.117794	Norway
5	Generosity	-0.083442	Norway
6	Perceptions of corruption	0.509541	Norway
7	Positive affect	0.640205	Norway
8	Negative affect	-0.446981	Norway

Figura 15: Tabla de correlación del país China con variable independiente Life Ladder y con el resto de variables del análisis

	Variable	Correlation	Country
0	Life Ladder	1.000000	China
1	Log GDP per capita	0.845717	China
2	Social support	0.453363	China
3	Healthy life expectancy at birth	0.839474	China
4	Freedom to make life choices	0.273188	China
5	Generosity	0.263581	China
6	Perceptions of corruption	NaN	China
7	Positive affect	0.268645	China
8	Negative affect	0.566617	China

Figura 16: Tabla de correlación del país Australia con variable independiente Life Ladder y con el resto de variables del análisis

	Variable	Correlation	Country
0	Life Ladder	1.000000	Australia
1	Log GDP per capita	-0.773033	Australia
2	Social support	0.630450	Australia
3	Healthy life expectancy at birth	-0.759425	Australia
4	Freedom to make life choices	0.775132	Australia
5	Generosity	0.380013	Australia
6	Perceptions of corruption	-0.644745	Australia
7	Positive affect	0.604250	Australia
8	Negative affect	-0.397398	Australia

Figura 17: Tabla de correlación del país Afghanistan con variable independiente Life Ladder y con el resto de variables del análisis

	Variable	Correlation	Country
0	Life Ladder	1.000000	Afghanistan
1	Log GDP per capita	-0.699539	Afghanistan
2	Social support	0.839457	Afghanistan
3	Healthy life expectancy at birth	-0.874994	Afghanistan
4	Freedom to make life choices	0.771925	Afghanistan
5	Generosity	0.656061	Afghanistan
6	Perceptions of corruption	-0.140289	Afghanistan
7	Positive affect	0.814103	Afghanistan
8	Negative affect	-0.829597	Afghanistan

Por otro lado, una vez vistas estas tablas de correlación por país queríamos plantear un modelo de regresión lineal que nos permitiera en términos generales para todos los países la fuerza que tienen esas correlaciones en las diferentes variables de la base de datos con la variable independiente de Life Ladder o índice de felicidad.

Figura 18: Coeficientes del modelo de regresión lineal por cada una de las variables teniendo como variable objetivo el Life Ladder

	Feature	Coefficient
0	Log GDP per capita	0.431933
6	Positive affect	0.247761
1	Social support	0.234755
2	Healthy life expectancy at birth	0.198985
5	Perceptions of corruption	-0.100265
3	Freedom to make life choices	0.070535
4	Generosity	0.054691
7	Negative affect	-0.017803
Error cuadrático medio del modelo: 0.2893		

1.6.3. Análisis de las correlaciones por país

Primero de todo, Noruega hemos ido viendo que de forma evolutiva es de los países más felices de este WHR. En que se basa la felicidad de uno de los países más felices? En tener menos dinero y en las emociones positivas. A menor PIB per cápita, mayor es la felicidad de este país nórdico y a mayores emociones positivas vividas antes de hacer la encuesta mayor es el índice de felicidad. En este país no se cumple que el dinero da la felicidad.

España también era interesante analizarla por nuestra cercanía y como con Noruega, el dinero a los españoles no nos da la felicidad. En España, de forma histórica desde 2006 hasta 2023 una variable que tiene alta correlación con el índice de felicidad es la del apoyo social, es decir, cuanto mayor es el apoyo y la red social tejida por la sociedad española

mayor es la felicidad. Otra variable que tiene alta correlación en España con la felicidad es la capacidad de poder tener libertad a la hora de tomar decisiones con un 0,67 de correlación. No obstante, con algo de lógica aunque veremos que no es así en todos los países, lo que hace más felices a los españoles es no haber vivido emociones negativas antes de hacer el cuestionario del WHR con una correlación de -0,72, es decir, con una correlación negativa.

Uno de los países más ricos se ha detectado que era Luxemburgo de forma histórica con el PIB per cápita más alto. Si se analiza su tabla de correlaciones con la variable Life Ladder la variable que tiene mayor correlación y, por tanto, más efecto es la de generosidad pero en negativo. Dicho de otra forma, Luxemburgo es un país que cuanto menos generoso más feliz es. Puede tener lógica teniendo en cuenta con el PIB per cápita alto que tienen. De hecho, la segunda variable con mayor correlación es la del PIB per cápita así que en Luxemburgo el dinero sí es importante y también lo es no hacer donaciones o no ser generoso como país.

Por otro lado, como representante de un continente como Oceanía tenemos Australia. Como el continente tiene dos variables muy interesantes que afectan a la felicidad y, además es más interesante que sea una correlación negativa. Se trata de un -0,77 en la variable del PIB per cápita y de un -0,75 en la esperanza de vida. Dicho de otra forma, los australianos cuanto menos viven y cuanto menos dinero tienen más felices son. Son dos variables que juntas muestran cómo es un país como Australia. Australia también tiene una alta correlación de la felicidad con la libertad de tomar decisiones y en este caso es positiva con un 0,77 siendo una de las más altas. Algo parecido que pasaba en España, es un país que valoran para ser felices el hecho de sentirse libres a la hora de tomar decisiones.

Por último dos países que serán opuestos en análisis. Se trata de China y Estados Unidos. Ambos tienen una correlación alta con la variable del PIB per cápita pero China tiene una correlación positiva de 0,84 y Estados Unidos tiene una correlación negativa de -0,78. De forma opuesta a unos les hace felices tener mayor PIB per cápita y a otros les hace menos felices tener mayor PIB per cápita. Si en algo están de acuerdo ambos países a la hora de ser felices es en la longevidad en sus vidas. Cuanto mayor es la esperanza de vida mayor es la felicidad de estos países.

Como consecuencia también del análisis realizado en este EDA también era de interés saber que pasaba con uno de los países que menos felices manifiesta ser. Se trata de Afghanistan que en los últimos años estaba en flop5 de los países menos felices. En este caso viendo la figura 17 se puede observar como la esperanza de edad y el PIB son variables con mayor correlación pero en negativo. Es decir, algo parecido que pasaba en China y Estados Unidos, como menos esperanza de vida y menor PIB per cápita, más felices son los habitantes de Afghanistan. No obstante, de todos los países analizados, el

país que más valora el apoyo de los familiares y amigos en el índice de felicidad es Afghanistan con un 0,82 de correlación de esa variable con el índice de felicidad.

Añadir que el hecho de que en China salgan nulos en la variable de percepción de corrupción es que esta pregunta no se debe realizar en China debido a situaciones políticas y en ningún año de los analizados se tienen datos. Como en este EDA se han tratado los nulos con la media, al no tener datos ningún año, siguen siendo valores nulos en este país.

En términos globales y planteando un estadístico que nos puede ayudar mucho a ver las relaciones que tiene las variables de esta EDA es la regresión lineal múltiple. Para ello se ha creado una función que nos ofrece el modelo estadístico que nos da esa regresión lineal de forma global, es decir, de todas las variables con la variable objetivo (life ladder o índice de felicidad) en todos los años y todos los países unidos (figura 18).

Interesa saber de esta regresión lineal ver la fuerza de las relaciones de todas las variables con la variable objetivo pero también el error cuadrático del modelo.

En primer lugar, se observa en la figura 18 como el PIB per cápita es la variable con mayor relación (0,43) con la variable de índice de felicidad. Esto se ha visto que no pasa de forma igual en todos los países, pero de forma global y para toda la base de datos, es la variable con mayor efecto en el WHR para ser feliz. En ese sentido, se cumple que el dinero da la felicidad.

Otras variables que tienen efecto de forma global en ser felices las hemos visto con correlación fuerte en algunos de los países analizados. Se trata de las variables de apoyo social y de emociones positivas antes de hacer el cuestionario. Son dos variables que dentro el modelo de regresión lineal y, de forma global, ofrecen una relación de 0,23 y 0,24.

El resto de variables tienen una relación muy cercana a 0 lo que hace afirmar que tienen menos fuerza a la hora de que la variable objetivo varíe.

Finalmente, mencionar que el MSE o error cuadrático medio del modelo que se ha extraído con la función es de 0,28. Un MSE (Mean Squared Error) de 0.28 indica el promedio de los cuadrados de las diferencias entre los valores reales y los valores predichos por el modelo. Este valor proporciona una medida de la precisión del modelo de regresión lineal. Cómo de preciso es este modelo entonces teniendo un valor de MSE de 0,28? Se trata de un valor relativamente bajo ya que está cercano a 0 (modelo ajustado perfecto) y, por lo tanto, el modelo de la función creada para esta EDA se ajusta bien a los datos.

2. Conclusiones

En este apartado se cerrarán todas las ideas de este EDA basado en todo el análisis realizado. Además se utilizarán las hipótesis planteadas en el apartado 1.2. Para ser mucho más claro en estas conclusiones se irá poniendo hipótesis a hipótesis y resolviendo cada

una de ellas con ideas muy esclarecedoras sobre la base de datos del WHR y de la EDA realizada.

1. Qué variables tienen un MAYOR efecto sobre el índice de felicidad o LIFE LADDER

Se ha detectado en esta EDA como veíamos en el análisis bivalente que la variable del PIB per cápita es la variable que mayor efecto tiene sobre el Life Ladder en términos globales. Otras variables que tienen efecto en la felicidad de los países es el apoyo social o confianza en los familiares y amigos y las emociones positivas. Dicho de otra forma, después del dinero, las redes sociales y el bienestar de las personas son otras variables a tener muy en cuenta para saber si un país es feliz o no.

2. Comparar diferentes países y la variable que tienen más efecto en el índice de felicidad o LIFE LADDER

Antes de realizar el EDA se planteaba si había diferencias en cómo se construye la felicidad según donde vives. Aunque puede parecer algo obvio, había que demostrarlo con datos con la EDA y así ha sido. Para ser feliz, no es lo mismo vivir en España que en Luxemburgo que en China que en Estados Unidos. De hecho, ya entre continentes existen diferencias ya que ,por ejemplo, en Oceanía la correlación el PIB per cápita es negativa contraria al resto de continentes.

En los países analizados también se han detectado diferencias ya que no en todos los países lo más importante para ser feliz era el PIB per cápita sino que en algunos como por ejemplo en España, el apoyo social era la variable que afectaba más en el índice de felicidad del país.

3. Las variables de generosidad y apoyo social han crecido desde la pandemia en el año 2020.

En el EDA se han podido ver con los gráficos extraídos que en el caso de la generosidad sí que subió en todos los continentes pero que no lo hizo de forma global el apoyo social. Hay matices dentro de esta hipótesis y es que se produce un efecto rebote con la variable de generosidad en la que 2021 sube en todos los continentes pero en la gran mayoría vuelve a bajar creando así la mencionada montaña de subida y bajada de la variable. Por su lado, el apoyo social pasa un poco el efecto inverso, es decir, disminuye en casi todos los continentes en el 2021 para subir en 2022 haciendo así el efecto rebote al revés.

4. Los países que lideran una variable no tienen porqué ser los más felices

Se ha podido analizar sobre todo la variable de riqueza y se han puesto de manifiesto los países más ricos de los últimos 6 años. Estos países más ricos (Luxemburgo, Irlanda) no aparecían en los países más felices año a año. Haciendo el análisis bivalente con las tablas de correlación se ha podido detectar que además en estos países ricos el PIB per cápita no es la variable con mayor efecto en el índice de felicidad sino que lo eran la

generosidad en el caso de Luxemburgo o la de apoyo social (y los pocos sentimientos negativos) en el caso de Irlanda.

5. Los 5 países MÁS felices del mundo tienen una ideología política basada en la socialdemocracia (busca apoyar las intervenciones estatales, tanto económicas como sociales, para promover mayor equidad económica e igualdad social en el marco de una economía capitalista)

Esta hipótesis tal vez era la más compleja de demostrar con datos pero podemos afirmar que, en su gran mayoría se ha cumplido. En el momento de extraer los 5 países más felices en los últimos 6 años (2017-2023) se ha podido ver que en su gran mayoría los países que aparecen en este top5 son Noruega, Finlandia, Dinamarca, Islandia o Suecia son países bajo una ideología política basada en un estado del bienestar socialdemócrata. No obstante, sorprende ver algún país con el que no se contaba como ser de los más felices. Es el caso de Costa Rica, Suiza o Israel los cuáles ahora especificaremos las principales ideas de sus estados del bienestar y veremos las diferencias con los países nórdicos (OECD Reports, 2024):

- Costa Rica: Fuerte énfasis en la educación y la salud pública, con un enfoque en la protección social.
- Israel: Estado de bienestar con servicios de salud y educación robustos, aunque con menor gasto público comparado con los países nórdicos.
- Países Nórdicos e Islandia: Modelos de bienestar socialdemócratas con altos niveles de gasto público y servicios sociales universales.
- Suiza: Economía de mercado con un sistema de seguridad social bien desarrollado, pero con menor gasto público que los países nórdicos.

De esta forma, todos los países exceptuando Suiza son países donde la apuesta por los servicios de salud y educación son fuertes aunque en los países nórdicos es mayor y con altos niveles de gasto público en esos servicios. Suiza es la única diferente con una economía de mercado pero que también tiene una seguridad social que ayuda a que las personas se sienten seguras en el país.

Por lo tanto, los países que más felices son en el WHR de los últimos 6 años (2017-2023) son países que apuestan por cuidar del ciudadano y apostar por gastar dinero público en sanidad y educación.

6. Los países mucho menos desarrollados económicamente (PIB más bajo) son los que tienen menor LIFE LADDER

En esta hipótesis se buscaba poner de manifiesto si realmente las personas más pobres pueden ser igualmente felices. Es cierto que cuando se han sacado los países más pobres durante todo el WHR desde 2005 hasta 2023 el índice de felicidad no supera en ningún

caso el 5. De hecho los 3 primeros países más pobres de toda la serie que es Venezuela (en los tres casos) en 2019, 2018 y 2017 tienen un índice de felicidad estable sobre el 5,0.

Cabe destacar, no obstante, que en esos años Venezuela vivió una hiperinflación super agresiva que hizo (junto a una devaluación del Bolívar y una caída en la producción petrolera) que se convirtiera en el país más pobre analizado en la serie del WHR. Entonces, si entramos a otros países que también aparecen en el top10 de países más pobres como Burundi, Congo (Kinshasa) o Central African Republic aquí el life ladder o índice de felicidad no supera el 3,5 y suspendería. En este sentido, existe una diferencia entre países menos desarrollados económicamente como los países africanos como el período muy concreto politicoeconómico (2017-2019).

Entonces, podemos afirmar que los países más pobres (y exceptuando el caso de Venezuela), son países donde el índice de felicidad es bajo.

Añadir solo a modo de cerrar esta hipótesis planteada que, países como Afganistán que aparecen en los países más infelices, no aparecen como los países más pobres por lo tanto el PIB per cápita muy bajo no es suficiente para ser el país más infeliz de la WHR.