

Documento base

Pablo Álvarez Arnedo

2026-02-11

Tabla de contenidos

1	Introducción	3
2	Datos longitudinales	4
2.1	Datos con medidas repetidas	4
2.2	Conceptos básicos de la regresión lineal simple	7
2.3	¿Por qué no se puede usar la estadística clásica?	8
2.3.1	Ejemplo conceptual	8
3	Modelos mixtos	15
3.1	Comparación de modelos con efectos fijos, aleatorios y mixtos	15
3.1.1	Modelo con efectos fijos	16
3.1.2	Modelo con efectos aleatorios	17
3.1.3	Modelo mixto	18
3.2	Modelos Lineales Mixtos (LMM)	19
3.3	Modelos Lineales Generalizados (GLM)	22
3.3.1	Ejemplo práctico	24
3.4	Modelos Lineales Generalizados Mixtos (GLMM)	25
3.4.1	Ejemplo práctico	26
4	Análisis de la base de datos	28
4.1	Análisis exploratorio inicial	31
4.2	Análisis de las bases de datos complementarias	32
4.3	Evolución de la felicidad a lo largo del tiempo	39
4.4	Evolución del Happiness Score en España	50
5	Construcción del modelo	65
5.1	Análisis exploratorio y selección inicial de variables	66
5.1.1	Estructura del dataset longitudinal	66
5.1.2	Visualización y evolución temporal de las variables	66
5.1.3	Matriz de correlaciones	70
5.2	Criterios de selección del modelo	71
5.3	Modelado clásico	72
5.3.1	Estrategia top-down (backward elimination)	72
5.3.2	Estrategia bottom-up (forward selection)	73
5.3.3	Diagnóstico y validación final del modelo	75

5.4	Modelos Lineales Mixtos (LMM)	77
5.4.1	Normalidad	83
5.4.2	Homocedasticidad	90
5.4.3	Outliers y estructura de los residuos	90
5.4.4	Conclusión del diagnóstico	90
5.4.5	Interpretación de coeficientes	90
5.5	Selección del mejor modelo	91
5.5.1	Normalidad	102
5.5.2	Homocedasticidad	102
5.5.3	Outliers y estructura de los residuos	102
5.5.4	Conclusión del diagnóstico	102
5.5.5	Interpretación de coeficientes	103
5.5.6	Predicción del Happiness Score para 2025	104
6	Aplicación Shiny para la modelización de la felicidad	106
6.1	Estructura general de la aplicación	106
6.1.1	Pestaña “Descriptiva”	106
6.1.2	Pestaña “Análisis”	107
6.1.3	Pestaña “Información”	107
6.2	Integración con el análisis longitudinal	107
6.3	Ventajas y potencial de la herramienta	108
6.4	Limitaciones y posibles mejoras	108
6.5	Conclusión	108
7	Conclusiones y mejoras futuras	109
7.1	Resumen del trabajo realizado	109
7.2	Valoración de la evolución del proyecto	110
7.3	Contribuciones del trabajo	110
7.4	Limitaciones	110
7.5	Posibles mejoras y extensiones futuras	111
7.6	Conclusión final	111
	Referencias	112

1 Introducción

```
1 + 1
```

```
[1] 2
```

2 Datos longitudinales

2.1 Datos con medidas repetidas

Los **datos longitudinales** son aquellos que obtenemos al realizar distintas medidas a un individuo (personas, regiones, células, etc.). Dichas medidas se pueden observar repetidamente a lo largo del tiempo (análisis temporal), como el ingreso anual de diferentes personas a lo largo de varios años; del espacio (análisis espacial), por ejemplo, al medir la contaminación del aire de distintas ciudades en un mismo día; o a lo largo del espacio y tiempo (análisis espacio-temporal), como puede ser la monitorización de la expansión de una enfermedad en distintas regiones a lo largo del tiempo. Como la causa más usual de medidas repetidas es el tiempo, haremos referencia a este caso en concreto, aunque los otros dos también serían aplicables. Por esto, a los datos longitudinales también se les conoce como medidas repetidas.

El análisis de este tipo de medidas nos permite detectar cambios o tendencias temporales en nuestras variables, lo cual nos puede llevar a observar patrones que nos sería difícil descubrir usando otro tipo de técnicas. Es común usar este tipo de datos en estudios donde se busca evaluar cómo evolucionan ciertas características o mediciones bajo distintas condiciones o tratamientos. En el ámbito biosanitario, los datos longitudinales son fundamentales para investigar la progresión de enfermedades, la efectividad de tratamientos y el impacto de intervenciones médicas. En este capítulo, exploraremos las características clave de los datos longitudinales y profundizaremos en las razones por las que los métodos clásicos, como la regresión lineal simple, fallan al aplicarse a este tipo de datos.

Como ya hemos mencionado anteriormente, una de las características que definen a los datos longitudinales es que tenemos medidas repetidas del mismo sujeto a través de diferentes observaciones. No obstante, dichas observaciones no están organizadas de cualquier manera, sino que están agrupadas por unidades (e.g., pacientes, regiones). Todo ello significa que cada unidad tiene varias observaciones en diferentes momentos temporales; haciendo que los datos longitudinales adopten una estructura jerárquica.

Esta estructura nos lleva a asumir una de las claves en todo este proceso, la dependencia entre las observaciones, la cual nos indica que las mediciones dentro de la misma unidad tienden a estar correlacionadas. También tenemos que destacar las distintas variables que definen a dichos datos, que suelen clasificarse según diferentes propiedades. Como la mayoría de medidas se realizan en distintos instantes de tiempo, es normal que su valor varíe a lo largo del tiempo, permitiendo considerarlas como variables tiempo-dependientes, lo que significa que sus cambios pueden estar relacionados con el tiempo y pueden ser modeladas para entender tendencias o

patrones; pero también hay que tener en cuenta que hay otras variables que cambian igual en el tiempo para todos los sujetos (como la edad) que no consideraremos tiempo-dependientes y otras que directamente son constantes como el sexo.

El análisis de datos longitudinales se centra en aprovechar las medidas repetidas para abordar preguntas específicas que no pueden ser respondidas adecuadamente con otros tipos de datos. Uno de los principales objetivos del análisis de estos datos es observar la evolución de una variable a lo largo del tiempo, lo cual nos permitiría poder detectar si los cambios de las variables siguen ciertos patrones o fluctuaciones que tendríamos que tener en cuenta en el análisis. Esta identificación de patrones nos puede aportar información y conocimientos clave, ya que nos ayuda a formular ciertas hipótesis que nos orientan hacia una visión concreta. Otra parte importante reside en comparar si la evolución de una variable a lo largo del tiempo es igual para distintas partes de la población, y ver si existen factores que regulan la evolución de dicha variable; en cuyo caso deberíamos de estudiar cómo dichos factores interactúan con el tiempo.

Los datos longitudinales tienen aplicaciones en una gran diversidad de áreas, ya que el estudio de medidas a lo largo del tiempo está presente en diferentes ámbitos científicos. Por ejemplo, los datos longitudinales tienen una gran importancia en el ámbito biosanitario, como puede ser en estudios donde hay medidas repetidas de presión arterial en un grupo de pacientes durante un tratamiento que nos permiten monitorear la salud de los pacientes para poder evaluar la efectividad del tratamiento e identificar posibles efectos secundarios. No obstante, este tipo de datos también tiene su relevancia en otras áreas como la educación; por ejemplo, la evaluación de los puntajes de un estudiante a lo largo de varios exámenes anuales podría identificar áreas de mejora por parte del alumnado o algunas estrategias pedagógicas que se puedan implementar en la docencia. Otra de las áreas en la que los datos longitudinales juegan un papel clave es la alimentación mediante el estudio de diferentes dietas a diferentes grupos de la población a lo largo del tiempo a través de medidas tales como la actividad física, peso corporal, nivel de colesterol, etc. y cómo estas rutinas aportan ciertos beneficios o riesgos a la salud de los individuos. En otros ámbitos como en el marketing también nos encontramos con casos en los que se utilizan datos longitudinales, como son encuestas de opinión realizadas periódicamente a las mismas personas que pueden ser de gran utilidad a la hora de evaluar posibles campañas de concienciación, o simplemente estudiar el comportamiento y la opinión de la población. Además, los datos longitudinales juegan un papel clave en el estudio de aspectos sociales, políticos y demográficos. Un ejemplo es el análisis de la felicidad y bienestar de los países a lo largo del tiempo, lo que permite identificar cómo factores como el crecimiento económico, la percepción de la corrupción, el acceso a servicios de salud y la cohesión social influyen en el bienestar de la población. Estos estudios pueden ser fundamentales para que los gobiernos diseñen políticas públicas que promuevan un mayor nivel de calidad de vida y bienestar social. También, en el ámbito demográfico, los datos longitudinales pueden ayudar a analizar la evolución de indicadores clave como la esperanza de vida, la migración o el desarrollo humano en diferentes regiones del mundo, proporcionando información valiosa para la toma de decisiones a nivel global.

A pesar de su gran utilidad, los datos longitudinales presentan varias complicaciones adicionales. En primer lugar, aunque las mediciones suelen realizarse en intervalos de tiempo predefinidos, no siempre disponemos de todas las observaciones esperadas debido a la presencia de valores faltantes. Esto puede ocurrir por razones como la ausencia de un paciente en una consulta médica, la falta de respuesta en una encuesta periódica o errores en la recolección de datos. Además, en muchos estudios, los individuos no necesariamente son medidos en los mismos instantes de tiempo, por lo que no siempre tenemos el mismo número de mediciones repetidas por individuo, lo que lleva a una estructura desigual en los datos que debe ser abordada con técnicas adecuadas. Estas dificultades pueden generar desafíos en el modelado y en la comparación de trayectorias individuales, por lo que es fundamental aplicar estrategias estadísticas como imputación de valores faltantes, modelado con efectos aleatorios o técnicas específicas para datos desbalanceados. Según Isaac Subirana en su *Curso de datos longitudinales* (Subirana 2020), los modelos lineales mixtos proporcionan una herramienta útil para abordar estos problemas, permitiendo modelar la estructura de correlación y manejar la heterogeneidad de las observaciones. Esto se puede apreciar en el siguiente ejemplo:

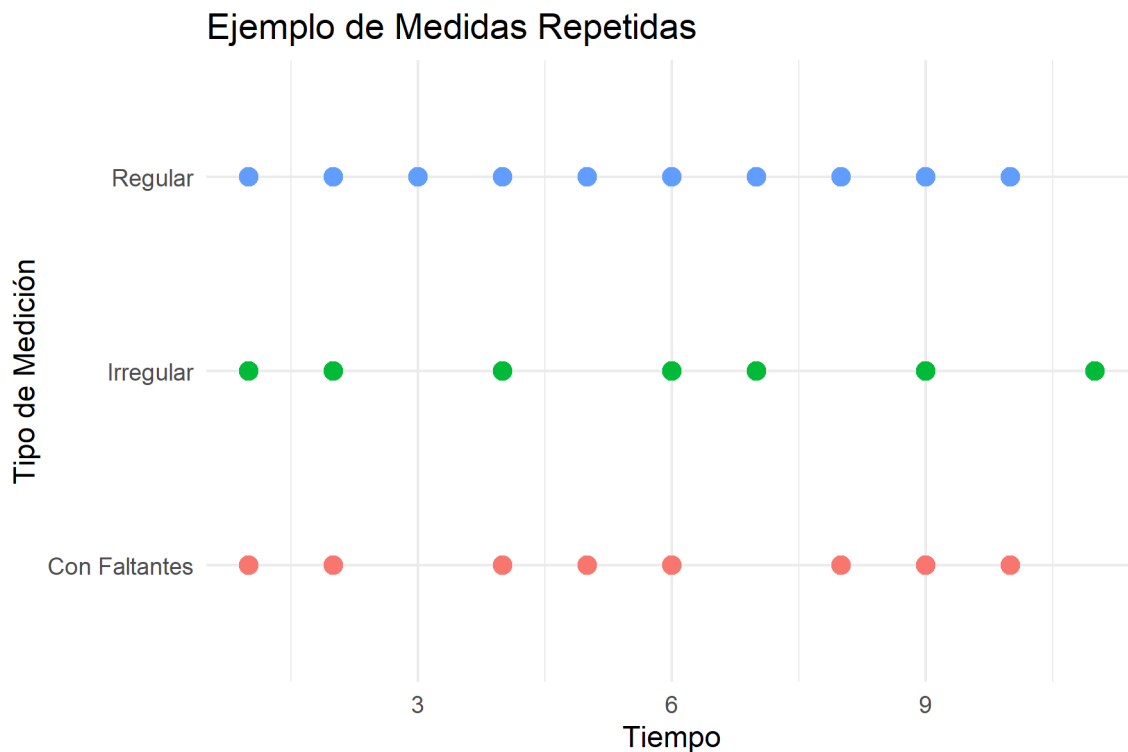


Figura 2.1: Ejemplo de medidas repetidas en diferentes estructuras temporales

Como podemos apreciar en la Figura 2.1, tenemos por un lado intervalos regulares en los que las mediciones se toman a intervalos de tiempo predefinidos, intervalos irregulares con valores ausentes en los que se han perdido algunas mediciones a lo largo del tiempo, y, por

último, intervalos irregulares en los que las mediciones no siguen una periodicidad fija. Estas complicaciones pueden suponer un problema, y es importante tenerlas en cuenta.

2.2 Conceptos básicos de la regresión lineal simple

La **regresión lineal simple** es un método estadístico utilizado para modelar la relación entre una **variable dependiente** Y (respuesta) y una **variable independiente** X (predictora) mediante una ecuación lineal. El modelo se define matemáticamente de la siguiente manera:

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$$

donde:

- Y_i representa la variable dependiente (respuesta).
- X_i es la variable independiente (predictora).
- β_0 es el **intercepto**, que indica el valor esperado de Y cuando $X = 0$.
- β_1 es la **pendiente**, que mide el cambio esperado en Y por cada unidad de cambio en X .
- ε_i representa el **término de error**, que captura la variabilidad no explicada por el modelo.

Para que la regresión lineal simple sea válida y produzca estimaciones confiables, deben cumplirse ciertos **supuestos** fundamentales:

1. **Linealidad:** La relación entre la variable independiente X y la dependiente Y debe ser lineal. Esto significa que un cambio en X se traduce en un cambio proporcional en Y .
2. **Independencia:** Las observaciones deben ser independientes entre sí. Es decir, los valores de Y no deben estar correlacionados con otras observaciones.
3. **Normalidad de los errores:** Se asume que los errores ε_i siguen una distribución normal con media cero ($\varepsilon_i \sim N(0, \sigma^2)$). Esto es especialmente importante para hacer inferencias estadísticas sobre los coeficientes β_0 y β_1 .
4. **Homocedasticidad:** La varianza de los errores debe ser constante para todos los valores de X . Es decir, la dispersión de los valores de Y en torno a la línea de regresión debe ser uniforme.

Cuando estos supuestos se cumplen, la regresión lineal simple proporciona **estimaciones insesgadas** de los coeficientes y permite hacer inferencia sobre la relación entre X y Y mediante pruebas de hipótesis y construcción de intervalos de confianza.

2.3 ¿Por qué no se puede usar la estadística clásica?

La estadística clásica (e.g., regresión lineal simple) parte de la suposición fundamental de que todas las observaciones son independientes entre sí. Sin embargo, en datos longitudinales, esta independencia no se cumple debido a la correlación entre mediciones repetidas de la misma unidad a lo largo del tiempo. Los longitudinales presentan características específicas que requieren enfoques estadísticos más avanzados.

Uno de los principales desafíos, ya mencionado anteriormente, es la dependencia entre observaciones, ya que los datos recogidos de un mismo individuo suelen estar correlacionados, lo que genera un patrón estructurado que no es capturado por modelos clásicos. Esta correlación también afecta a la estructura de los errores, ya que las mediciones repetidas pueden estar influenciadas por factores externos o por variables no observadas, lo que genera una relación entre los errores que los modelos tradicionales no pueden modelar correctamente. Además, la variabilidad entre individuos es un aspecto clave en datos longitudinales, ya que no todos los sujetos presentan la misma evolución en el tiempo. Los modelos clásicos suelen asumir una varianza homogénea, lo que no es adecuado en este contexto, ya que no permite capturar diferencias individuales ni estructuras de correlación complejas.

Todos estos factores hacen que el uso de modelos estadísticos convencionales, como la regresión lineal simple, no sea adecuado para el análisis de datos longitudinales. En su lugar, es necesario recurrir a enfoques específicos, como los modelos lineales mixtos, que permiten modelar tanto los efectos fijos como los efectos aleatorios para capturar adecuadamente la variabilidad y dependencia inherente a estos datos. La mejor manera de comprender estas limitaciones es a través de un ejemplo práctico.

2.3.1 Ejemplo conceptual

Para ilustrar las limitaciones de la estadística clásica en el análisis de datos longitudinales, vamos a considerar un conjunto de datos sobre ingresos anuales (en euros) de 10 personas medidos a lo largo de varios años (*psid*). Vamos a utilizar un modelo regresión lineal simple para modelar los ingresos en función del tiempo, ignorando la correlación entre mediciones.

En este ejemplo, la variable **dependiente** Y es el **ingreso anual** de cada persona; mientras que la variable **independiente** X es el **año**, representando el tiempo.

El objetivo del modelo es analizar si existe una tendencia en la evolución de los ingresos y, en caso afirmativo, estimar la relación entre el año y el nivel de ingresos de los individuos. Sin embargo, al aplicar un modelo de regresión lineal simple, ignoraremos la dependencia entre las observaciones de cada persona, lo que resultará en una estimación sesgada y poco fiable.

La Figura 2.2 muestra la evolución de los ingresos anuales para diferentes personas a lo largo del tiempo, en el que cada línea representa a una persona.

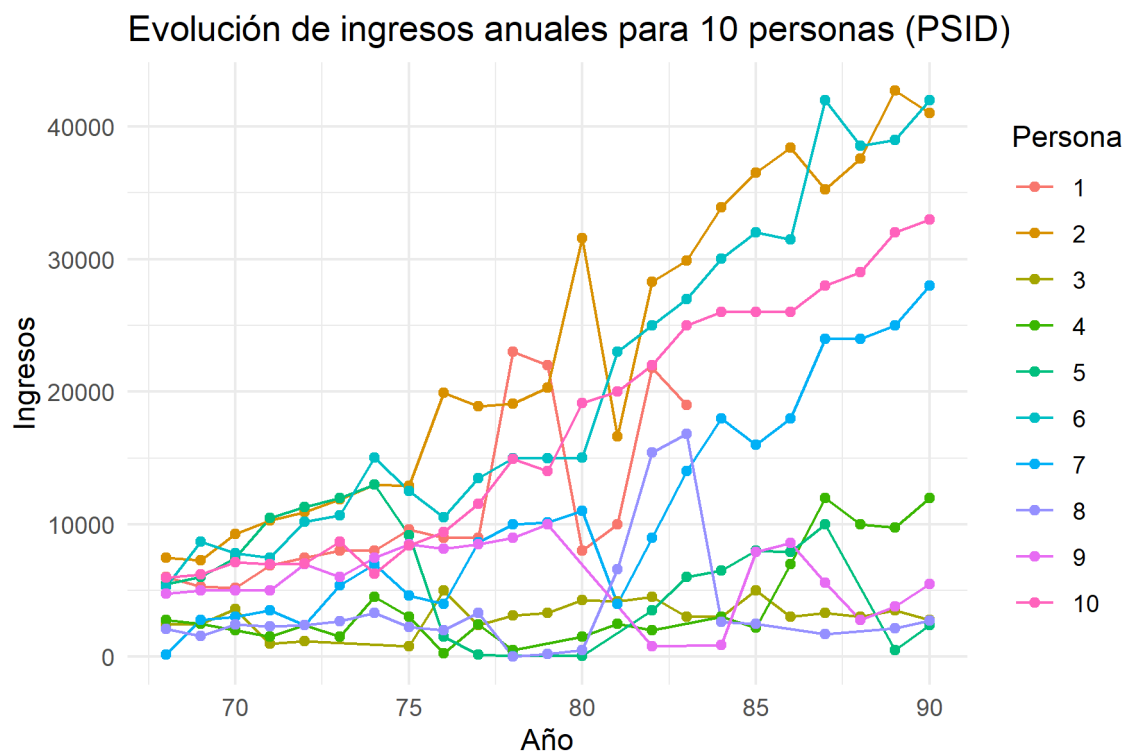


Figura 2.2: Evolución de los ingresos anuales de 10 personas a lo largo del tiempo

Esto permite mostrar cómo los ingresos varían entre individuos y años, observando que los datos son heterogéneos y varían significativamente entre individuos. Sin embargo, dentro de cada individuo, los ingresos en un año determinado tienden a ser similares a los del año anterior y el siguiente, lo que sugiere una correlación temporal en las mediciones. Esta dependencia entre observaciones dentro de cada individuo es una característica fundamental de los datos longitudinales, ya que implica que el valor de la variable en un momento dado está influenciado por valores previos del mismo individuo; algo que viola los supuestos básicos de independencia de las observaciones, fundamentales para modelos clásicos como la regresión lineal simple.

Visto esto, modelaremos la relación entre los ingresos y el tiempo utilizando una regresión lineal simple, ignorando la dependencia entre observaciones, para mostrar las consecuencias de no cumplir las hipótesis requeridas. La Figura 5.5 muestra el ajuste de la regresión lineal simple aplicada a los datos.

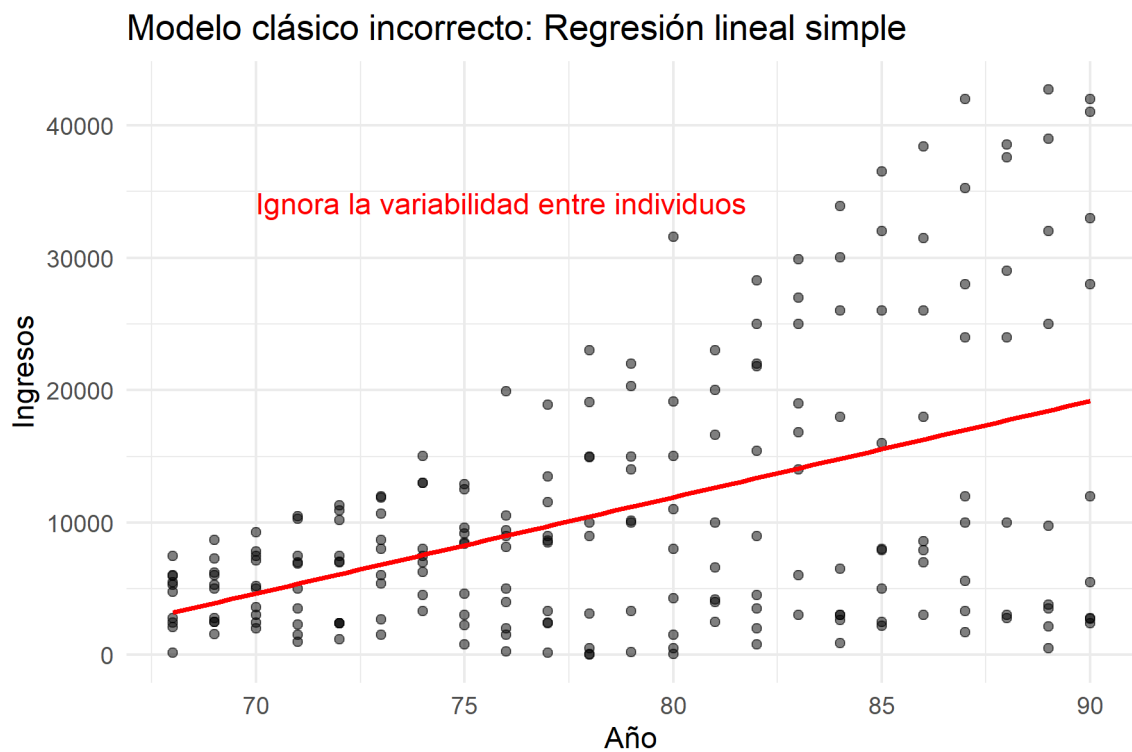


Figura 2.3: Ajuste del modelo de regresión lineal simple ignorando estructura longitudinal

La Figura 5.5 muestra cómo la regresión lineal simple aplicada a estos datos genera una representación distorsionada, ignorando por completo la correlación de los datos longitudinales; dando lugar a un mal ajuste y a resultados estadísticos inapropiados que demuestran por qué no debemos utilizar estadística clásica para este tipo de datos. No obstante, vamos a analizar la adecuación y diagnóstico del modelo para ver en detalle los motivos por los que las técnicas de estadística clásica no son las correctas para trabajar con datos longitudinales.

Al utilizar un modelo de regresión lineal simple estamos asumiendo que la variabilidad entre individuos se puede representar con un único coeficiente, ignorando por completo la dependencia entre observaciones. Para evaluar la adecuación del modelo, nos fijamos en una medida de bondad de ajuste como el coeficiente de determinación R^2 . El R^2 obtenido (**0.217**) es muy bajo, indicando que el modelo explica muy poca variabilidad en los datos (21%) y que, por tanto, no nos sirve para analizar datos longitudinales ya que no captura adecuadamente la relación entre las variables.

Para realizar el diagnóstico del modelo, haremos un análisis de los residuos. Recordemos que dicho análisis se basa en 4 partes fundamentales: la normalidad de los residuos, que tengan media cero, la no correlación y la homocedasticidad.

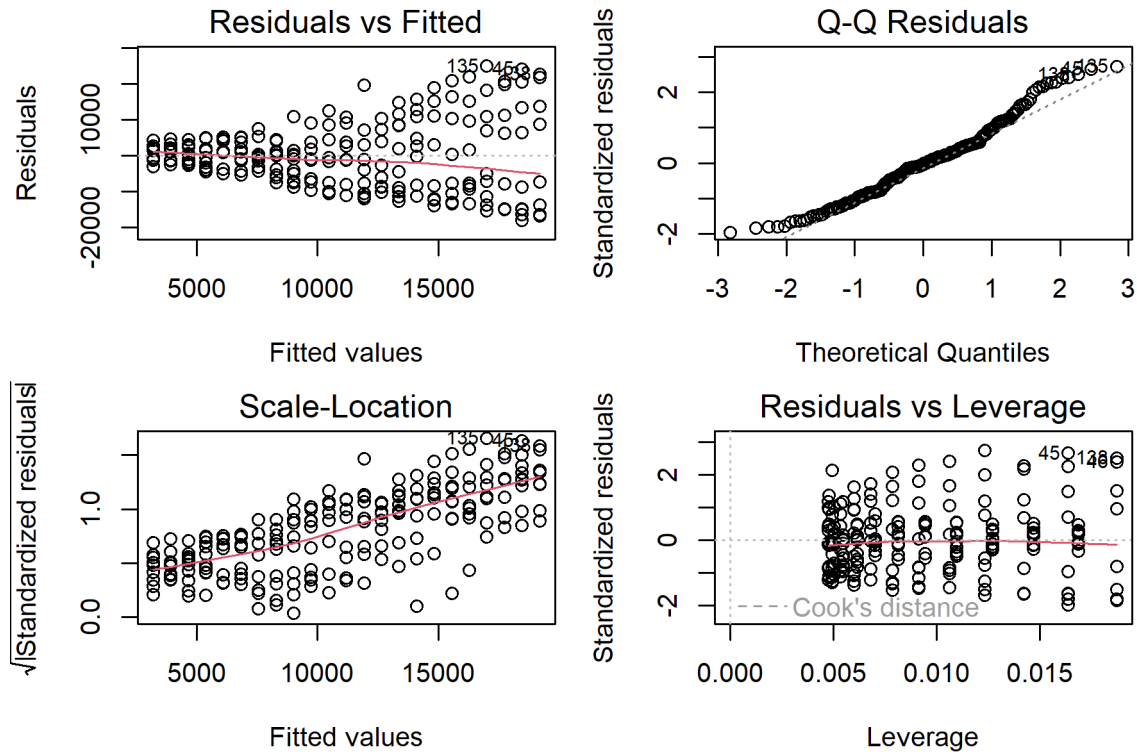


Figura 2.4: Gráfica de los residuos del modelo

Primero de todo, vamos a analizar es la **media cero** de los residuos. Su hipótesis de asunción es la siguiente:

$$\begin{cases} H_0 : \text{Los residuos tienen una media esperada de 0.} \\ H_1 : \text{Los residuos no tienen una media esperada de 0.} \end{cases}$$

Si calculamos la media de los residuos del modelo, comprobamos que la media es **0**, pero esta no es una forma correcta de analizar la media cero ya que esto no significa que la suposición

de media cero se cumpla en todas partes del rango de los valores ajustados. Para hacer un correcto análisis, nos vamos a fijar en la primera gráfica de la Figura 2.4: Residuals vs Fitted. Teóricamente, para que los residuos tengan media cero, deberían de estar uniformemente dispersos alrededor del eje horizontal en $y = 0$. Viendo la gráfica, podemos observar que los errores no tienen media cero ya que para los valores ajustados más altos se alejan mucho de la recta $y = 0$; por lo que esta es otra muestra más de que el modelo no es correcto para este tipo de datos.

Lo segundo que vamos a analizar es la **no correlación** entre los errores, la cual se puede analizar en la primera gráfica. Si nos fijamos, se observa un patrón curvilíneo a medida que aumenta el valor de los valores ajustados, por lo que se podría concluir que los errores están correlacionados. No obstante, para una verificación numérica haremos un test de Durbin-Watson para comprobar la no correlación. El test de Durbin-Watson verifica si los residuos están correlacionados en el tiempo. Sus hipótesis son las siguientes:

$$\begin{cases} H_0 : \text{No hay autocorrelación entre los residuos.} \\ H_1 : \text{Existe autocorrelación entre los residuos.} \end{cases}$$

En efecto, haciendo el test de Durbin-Watson vemos como el p-valor (**0**) es extremadamente bajo y nos permite concluir que podemos rechazar la hipótesis nula. Por tanto, podemos asumir que la correlación entre los errores no es 0; otro motivo más para ver que este modelo no funciona bien con datos longitudinales.

La tercera parte que vamos a analizar es la **normalidad** de los residuos. Para ello, nos fijamos en la gráfica superior derecha (Normal Q-Q) de la Figura 2.4, en la cual vemos que, aunque la mayoría de los puntos se alinean con la línea teórica, no son pocas las desviaciones que hay en los extremos; lo que sugiere que los residuos no son perfectamente normales. De hecho, también puede ser el caso paradigmático de normalidad heterocedástica, en la que la varianza depende de la media. Para salir de dudas, podemos aplicar un test de Jarque Bera. El test de Jarque Bera comprueba si los residuos siguen una distribución normal evaluando su asimetría y curtosis. Sus hipótesis son las siguientes:

$$\begin{cases} H_0 : \text{Los residuos siguen una distribución normal.} \\ H_1 : \text{Los residuos no siguen una distribución normal.} \end{cases}$$

Si el p-valor es menor a un umbral significativo (por defecto decimos que es 0.05), se rechaza la hipótesis nula, indicando que los residuos no siguen una distribución normal.

A través de este test, el p-valor (**0.024**) nos permite concluir que podemos rechazar la hipótesis nula y que, por tanto, los residuos no tienen normalidad.

Por último, analizaremos la **homocedasticidad** de los errores. Para ello, nos fijaremos en la primera (Residuals vs Fitted) y en la tercera gráfica (Scale-Location). A través de la gráfica Residuals vs Fitted, vemos como los residuos no tienen una varianza constante, sino que a medida que aumenta el valor de los valores ajustados aumenta su dispersión; por lo que no

tienen homocedasticidad, sino heterocedasticidad. Mirando la gráfica Scale-Location, podemos observar una tendencia creciente por parte de los residuos que nos permite ver cómo no tienen varianza constante. Para confirmarlo, haremos un test de Breusch-Pagan. El test de Breusch-Pagan evalúa si los residuos presentan heterocedasticidad; es decir, si su varianza no es constante. Sus hipótesis son las siguientes:

$$\begin{cases} H_0 : \text{Los residuos tienen varianza constante (homocedasticidad).} \\ H_1 : \text{Los residuos no tienen varianza constante (heterocedasticidad).} \end{cases}$$

De nuevo, vemos cómo el p-valor (<0.001) es extremadamente pequeño, lo que nos permite rechazar la hipótesis nula y, por lo tanto, concluir que los residuos no tienen varianza constante.

A través de este análisis, hemos podido comprobar que no podemos usar modelos de estadística clásica, tal y como la regresión lineal simple, para trabajar con datos longitudinales.

Una visión más acertada sería utilizar un modelo que se ajuste a cada individuo, como se hace en la Figura 2.5.

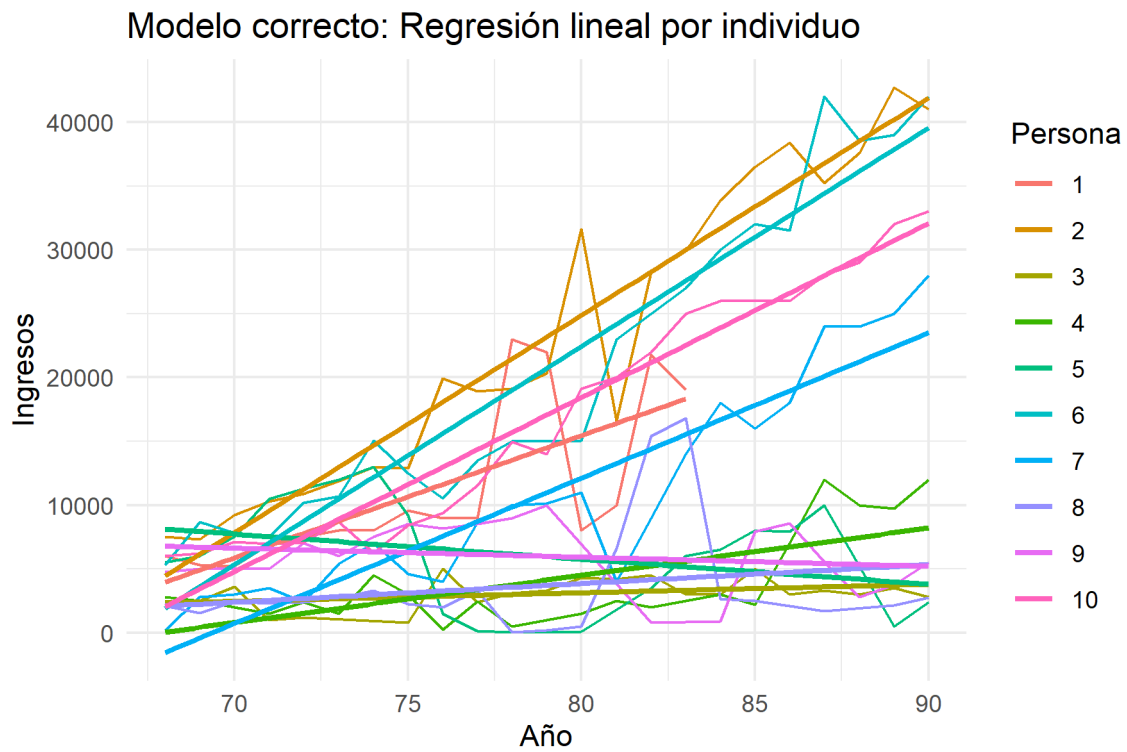


Figura 2.5: Ajuste de un modelo lineal individualizado para cada sujeto

En esta Figura 2.5, podemos observar que cada individuo tiene un comportamiento único en cuanto a la evolución de sus ingresos a lo largo del tiempo. Los interceptos y las pendientes varían considerablemente entre las personas, lo que evidencia que un único modelo no puede capturar adecuadamente la relación entre el tiempo y los ingresos para todos los individuos. Este resultado destaca la heterogeneidad presente en los datos y la necesidad de utilizar modelos que consideren esta variabilidad. Al ajustar un modelo por cada individuo, capturamos mejor las características específicas de cada sujeto, pero esta estrategia presenta limitaciones: aunque mejora la representación de la variabilidad entre individuos, no permite hacer inferencias generales sobre la población; además de que en escenarios con un gran número de individuos, esta aproximación no es práctica. Por ello, los **modelos mixtos**, que se explicarán en el siguiente capítulo, emergen como una solución adecuada, ya que combinan los llamados efectos fijos y aleatorios para capturar tanto las tendencias generales de la población como las diferencias específicas entre individuos. Esta aproximación ofrece un equilibrio entre flexibilidad y generalización, respetando las características únicas de los datos longitudinales.

3 Modelos mixtos

En este capítulo, exploraremos los **Modelos Lineales Mixtos (LMM)** y los **Modelos Lineales Generalizados (GLM)**, dos enfoques estadísticos fundamentales para el análisis de datos longitudinales. Veremos cómo los LMM permiten modelar la variabilidad entre individuos mediante la inclusión de efectos aleatorios y fijos, lo que facilita el estudio de la correlación entre observaciones repetidas. Luego, introduciremos los GLM, que extienden la regresión lineal para manejar variables respuesta que no siguen una distribución normal, utilizando funciones de enlace y la familia exponencial. A lo largo del capítulo, revisaremos sus formulaciones matemáticas, sus hipótesis clave y cómo validarlas en la práctica.

Para ilustrar estos modelos, comenzaremos con un ejemplo aplicado al conjunto de datos Orthodont del paquete nlme, donde analizaremos la evolución de la distancia entre los dientes (distance) en función de la edad (age) en diferentes sujetos. Compararemos tres enfoques distintos:

- **Modelo con sólo efectos fijos:** Se asume que todos los sujetos siguen la misma relación.
- **Modelo con sólo efectos aleatorios:** Se permite que cada sujeto tenga su propio valor inicial (intercepto), pero no afecta la pendiente.
- **Modelo mixto:** Se permite que tanto el intercepto como la pendiente varíen entre sujetos.

3.1 Comparación de modelos con efectos fijos, aleatorios y mixtos

La base de datos Orthodont proviene del paquete nlme en R y contiene información sobre el crecimiento dental en niños. Sus variables principales son:

- distance: distancia entre los dientes (variable respuesta).
- age: edad del niño (variable predictora principal).
- Subject: identificador del niño (variable de agrupación para efectos aleatorios).

A continuación, ajustaremos y visualizaremos los distintos modelos.

3.1.1 Modelo con efectos fijos

El primer modelo que consideramos es una regresión lineal simple, en la que asumimos que la distancia interdental (*distance*) varía en función de la edad (*age*), pero asumimos que todas las observaciones son independientes e ignoramos la estructura jerárquica del estudio (mediciones repetidas por individuo). La ecuación del modelo es:

$$distance_i = \beta_0 + \beta_1 age_i + \epsilon_i$$

Aquí, $distance_i$ es la distancia interdental de la observación i , β_0 es la intersección común a todos los sujetos, β_1 es la pendiente (cómo cambia la distancia con la edad), y ϵ_i es el error aleatorio.

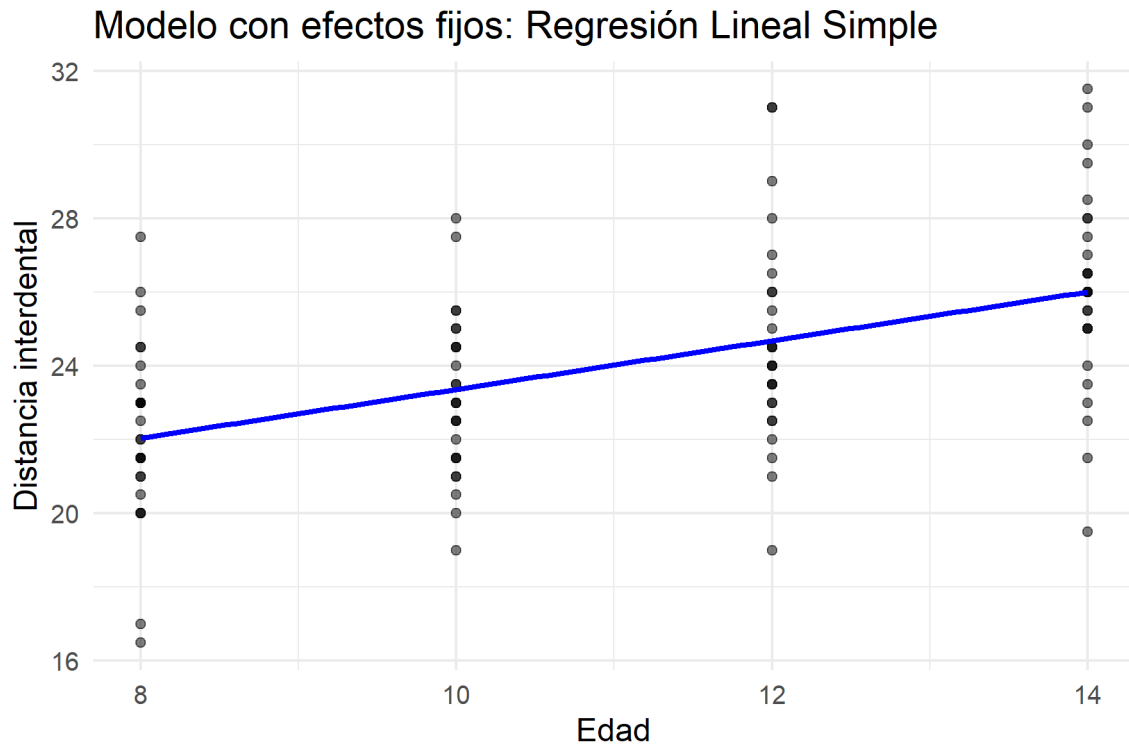


Figura 3.1: Modelo con efectos fijos: Regresión Lineal Simple

Este modelo de la Figura 3.1 considera únicamente la edad (*age*) como predictor de la distancia (*distance*) y no tiene en cuenta que los datos son mediciones repetidas de los mismos individuos, lo que puede llevar a errores de estimación debido a la correlación entre observaciones de un mismo sujeto. Como podemos comprobar a través de este ejemplo, si se ignora la estructura jerárquica, podríamos obtener estimaciones erróneas de la variabilidad en la población; obteniendo un coeficiente de determinación R^2 bajísimo (**0.256**).

3.1.2 Modelo con efectos aleatorios

Ahora ajustamos un modelo con efectos aleatorios, en el que permitimos que cada niño tenga su propio intercepto aleatorio (u_i), capturando la variabilidad entre individuos. La ecuación del modelo es:

$$distance_{ij} = \beta_0 + u_i + \beta_1 age_{ij} + \epsilon_{ij}$$

Ahora $distance_{ij}$ es la distancia para el individuo i en la observación j . u_i representa el efecto aleatorio de cada sujeto (intersección específica), β_1 es la pendiente común, y ϵ_{ij} es el error.

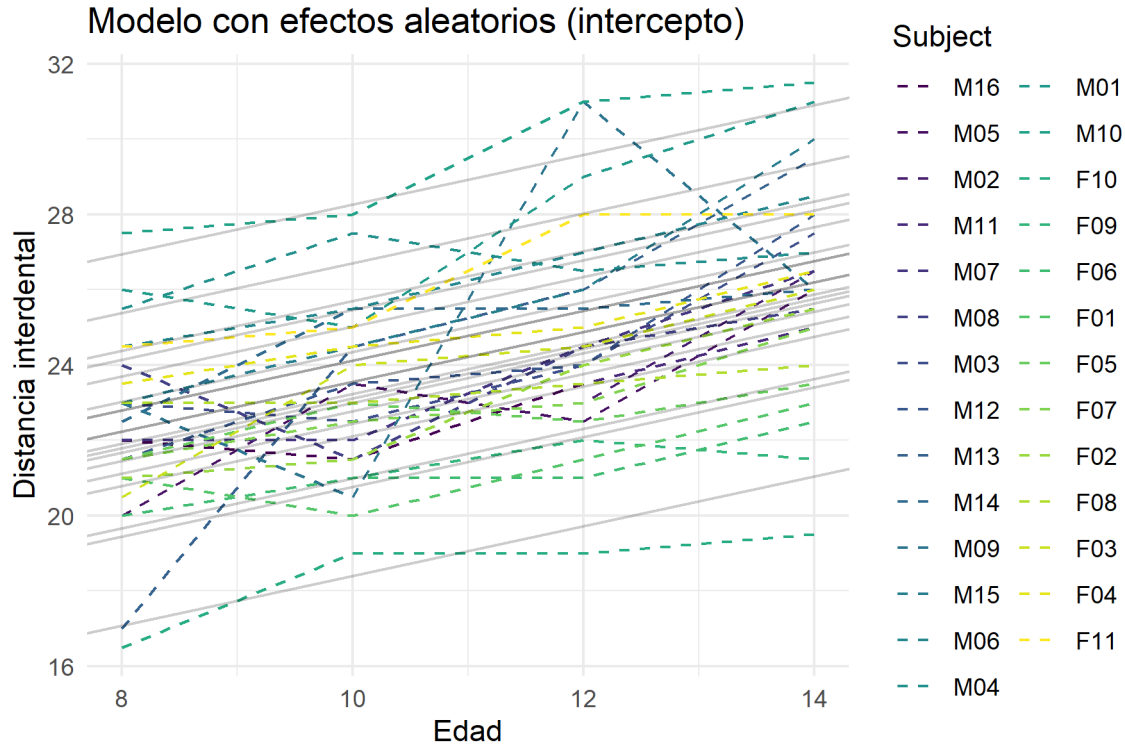


Figura 3.2: Modelo con efectos aleatorios a través de intercepto aleatorio

Como podemos apreciar en la Figura 3.2, ahora tenemos un término indica que cada individuo (Subject) tiene su propia intersección aleatoria; permitiendo que la relación entre la distancia y la edad varíe entre individuos en lugar de asumir una única intersección fija para todos. Esto significa que algunos sujetos pueden tener valores iniciales más altos o más bajos de distance sin que eso afecte la tendencia general de la población. La diferencia crucial de los efectos aleatorios la podemos apreciar en la variabilidad del modelo, ya que tenemos una varianza del intercepto por sujeto de **4.472** y una varianza residual de **2.049**, lo que significa que cada sujeto tiene un punto de partida diferente en distance, pero que todavía hay una parte de la variabilidad del modelo que no se explica por los efectos fijos ni por las diferencias entre sujetos.

Este modelo permite que cada niño tenga su propio intercepto aleatorio, modelando mejor la variabilidad individual.

3.1.3 Modelo mixto

Finalmente, ajustamos un **Modelo Lineal Mixto (LMM)** en el que consideramos tanto efectos fijos como aleatorios. Permitimos que cada niño tenga su propio intercepto (u_i) y pendiente (v_i) aleatorios, permitiendo que la relación entre edad y distancia interdental varíe entre individuos. La ecuación del modelo es:

$$distance_{ij} = \beta_0 + u_i + (\beta_1 + v_i)age_{ij} + \epsilon_{ij}$$

Aquí u_i es la intersección específica de cada sujeto, y v_i permite que la pendiente también varíe por individuo.

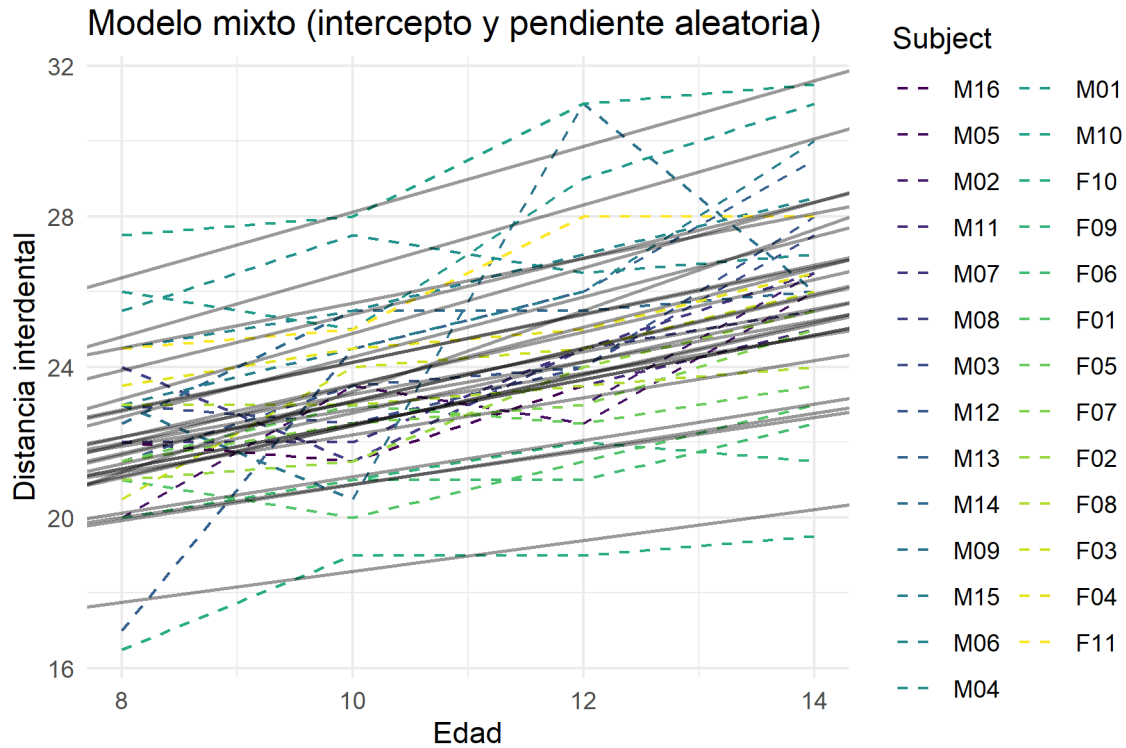


Figura 3.3: Modelo Lineal Mixto (LMM con efectos fijos y aleatorios)

Ahora no solo permitimos una intersección aleatoria, sino que también permitimos que la pendiente (efecto de la edad) varíe entre sujetos; es decir, en este modelo cada sujeto puede tener una tasa de crecimiento diferente en la distancia dental a lo largo del tiempo. Observando el modelo de la Figura 3.3, vemos como ahora hemos reducido la varianza residual a **1.716**,

y ahora contamos con una varianza del intercepto por sujeto de **5.417** y una variación de la pendiente entre sujetos de **0.051**; obteniendo una mejora significativa. Este tipo de modelos es más realista cuando hay variabilidad individual en la evolución de la variable respuesta.

Este modelo es más flexible, ya que permite que tanto la intersección como la pendiente de la relación entre edad y distancia varíen entre individuos. Este último modelo generaliza la idea que vimos en el capítulo anterior, donde ajustábamos una regresión por individuo (Figura 4). En aquel caso, teníamos una pendiente e intercepto diferentes por persona, pero ajustados de forma separada. Los modelos mixtos permiten hacer esto mismo, pero de forma conjunta y eficiente, combinando la información de todos los individuos para obtener estimaciones más robustas, sin necesidad de ajustar un modelo por separado para cada uno.

Si comparamos los 3 modelos, podemos observar que el modelo con solo **efectos fijos** asume una única relación entre edad y distancia interdental, ignorando la variabilidad entre individuos. El modelo con solo **efectos aleatorios** permite que cada sujeto tenga su propio intercepto, pero mantiene una pendiente común para todos. El **modelo mixto (LMM)** es el más completo, permitiendo que tanto la intersección como la pendiente varíen entre individuos. Esto demuestra la importancia de los Modelos Lineales Mixtos en el análisis de datos longitudinales, ya que incorporan tanto la variabilidad individual como la estructura jerárquica de los datos.

3.2 Modelos Lineales Mixtos (LMM)

Son métodos y modelos estadísticos que sirven para analizar datos longitudinales cuando la variable respuesta sigue una distribución normal. Uno de sus aspectos más característicos lo indica Francisco Hernández-Barrera en su libro *Modelos mixtos con R* (Hernández-Barrera 2024), ya que se asume que existe una relación entre el vector de observaciones y las covariables. Se considera la técnica más eficaz cuando se trabaja con distribuciones normales en este campo ya que permite introducir efectos aleatorios y concretar la estructura de las correlaciones de los residuos del mismo sujeto; además de que puede emplearse con datos faltantes. Estos modelos nos permiten modelar la correlación entre observaciones dentro de una misma unidad e incluir covariables tanto a nivel individual como grupal. Los LMM permiten realizar una estimación precisa de la incertidumbre, respetando la dependencia entre observaciones. Por otro lado, su capacidad de generalización a estructuras de datos complejas es otro de los motivos por los cuales se recomienda su uso con datos longitudinales. Otra de sus ventajas es su flexibilidad para incluir efectos específicos por individuo o grupo; algo que veremos más adelante.

La ecuación para este tipo de modelos, en los que y_{ij} representa el momento j -ésimo del individuo i :

$$y_{ij} = \beta_{0i} + \sum_{k=1}^K \beta_{ki} x_{ijk} + e_{ij}$$

- x_{ijk} es el valor de la k -ésima variable independiente por parte del individuo i en la observación j .

- β_{0i} sigue $N(\beta_0, \sigma_{\beta_0}^2)$; es el intercepto del modelo, que suele tener cierta varianza centrada en μ porque se supone aleatoria.
- β_{ki} sigue $N(\beta_k, \sigma_{\beta_k}^2)$; son las pendientes o coeficientes de las variables independientes del modelo, que suelen ser aleatorias.

Los **efectos aleatorios** se representan mediante el vector formado por la constante y los coeficientes aleatorios del modelo. Nos permiten capturar la variabilidad entre individuos, y se escriben de esta forma:

$$\vec{\beta}_i = (\beta_{0i}, \beta_{1i}, \dots, \beta_{Ki})^t \sim N(\vec{\beta}, \Omega)$$

Cabe destacar que los errores de un individuo, al no tener todos el mismo número de observaciones, son **independientes** de los efectos aleatorios.

Para ajustar un modelo lineal mixto, se tienen que disponer los datos de forma vertical. Una de las ventajas del LMM es su flexibilidad ya que no sólo permite especificar efectos aleatorios para evaluar la **variabilidad** de algunas variables entre los individuos, sino que también permite evaluar la **correlación** entre distintos datos longitudinales del mismo individuo. La constante y los coeficientes aleatorios tienen **homocedasticidad**, ya que la esperanza y la matriz de covarianzas es la misma para todos los individuos. Una de las características de los LMM es que introducen el concepto de **efectos fijos**, los cuales son la esperanza de los efectos aleatorios. Según Julian Faraway en *Extending the Linear Model with R: Generalized Linear, Mixed Effects and Nonparametric Regression Models* (Faraway 2006), un efecto fijo es una constante desconocida que intentaremos estimar a partir de los datos, mientras que los efectos aleatorios son variables aleatorias. De hecho, cuando un coeficiente no es aleatorio, se puede asumir que sigue una distribución normal con varianza cero; denominándolo fijo. En estos modelos, el número de efectos aleatorios es limitado, ya que no pueden superar en ningún caso el número de medidas que tenemos por individuo. Estos efectos inducen una **correlación** entre datos para el mismo individuo, pero dependiendo de su estructura la correlación sólo se puede obtener a partir de la correlación entre residuos; ya que si consideramos los coeficientes de variables **cambiantes** en el tiempo como aleatorios la correlación es distinta según los tiempos de las medidas, mientras que si consideramos los coeficientes de variables **constantes** en el tiempo inducimos **heterocedasticidad** entre individuos.

Una vez formulado el modelo, necesitamos estimar los parámetros, contando con los dos principales métodos de estimación en modelos mixtos. El primero es el método de Máxima verosimilitud (ML), que utiliza la función de verosimilitud completa del modelo. Estima tanto los efectos fijos como las varianzas de los efectos aleatorios. Es útil para comparar modelos con diferentes efectos fijos. El otro método es el de máxima verosimilitud restringida (REML), que estima solo las varianzas de los efectos aleatorios, ajustando los grados de libertad para evitar el sesgo en la estimación de la varianza. Es el método preferido para comparar modelos con la misma estructura de efectos fijos pero distinta estructura de efectos aleatorios.

A la hora de trabajar con Modelos Lineales Mixtos, se puede trabajar de diferentes formas. Podemos establecer un modelo con la constante aleatoria y varios coeficientes fijos en el tiempo,

en cuyo caso, si asumimos que los errores son independientes, tendríamos una correlación **constante** entre las variables del mismo individuo que no depende de la distancia entre las medidas; lo que se denomina como coeficiente de correlación intraclase (ICC). Otra forma de definir estos modelos podría ser con la constante y los coeficientes aleatorios, donde, asumiendo independencia entre residuos, la correlación entre observaciones pasa a depender tanto del tiempo como de la distancia entre ellas. Sin embargo, pese a ser las dos buenas opciones, es preferible trabajar de otra forma para LMM.

Para empezar, no asumiremos independencia de los residuos; sino que trabajaremos con un modelo más general en el que contemos con el mayor número posible de efectos aleatorios correlacionados y fijos. A continuación, procederemos a simplificar el modelo a través de la significación de **efectos aleatorios**:

$$\begin{cases} H_0 : \sigma_{\beta_0}^2 = 0 \\ H_1 : \sigma_{\beta_0}^2 > 0 \end{cases}$$

Para comprobar que hay más de un efecto aleatorio significativo, se utilizan diferentes técnicas estadísticas para contrastar que se permite asumir que podemos rechazar la hipótesis nula: los efectos aleatorios tienen varianza igual a cero. En caso afirmativo, tenemos que contrastar que si su correlación es distinta de 0; para lo que tendremos que elegir la matriz de covarianzas de los efectos aleatorios:

$$\begin{cases} H_0 : \Omega = \begin{pmatrix} \sigma_{\beta_0}^2 & 0 \\ 0 & \sigma_{\beta_1}^2 \end{pmatrix} \\ H_1 : \Omega = \begin{pmatrix} \sigma_{\beta_0}^2 & \sigma_{\beta_0\beta_1} \\ \sigma_{\beta_0\beta_1} & \sigma_{\beta_1}^2 \end{pmatrix} \end{cases}$$

En este caso, utilizaremos un test de razón de verosimilitudes para escoger la estructura de covarianzas de los efectos aleatorios y de sus errores. Para ello, hay que tener en cuenta que los modelos estén **anidados**, es decir, que la matriz de covarianzas de los residuos de un modelo se expresen como un caso particular de la matriz de covarianzas de los residuos del otro modelo.

Una vez hemos terminado con los efectos aleatorios, procedemos a determinar la significación de los efectos fijos a través de dos métodos. Si queremos testear un sólo parámetro, utilizaremos el test de Wald en el que:

$$\begin{cases} H_0 : \beta_1 = 0 \\ H_1 : \beta_1 \neq 0 \end{cases}$$

En caso de querer testear más de un parámetro, utilizaremos un test de razón de verosimilitudes:

$$\begin{cases} H_0 : \beta_1 = \beta_2 = 0 \\ H_1 : \text{alguno diferente de } 0 \end{cases}$$

Una vez hemos definido ya nuestro modelo, tenemos que realizar su validación a través de comprobar que se cumplen las asunciones sobre los residuos; al igual que hacíamos con Regresión Lineal Simple. Para poder asumir que el modelo es correcto, en el gráfico **residuos**

estandarizados vs valores predichos, debería de aparecer una especie de nube de puntos en los que no haya ningún patrón ni ninguna tendencia aparente; mientras que en el **QQ-plot**, si los residuos se encuentran alrededor de la diagonal sin seguir tampoco ningún patrón, podremos asumir que los residuos tienen normalidad. Para validar los efectos aleatorios, podemos utilizar **Empirical Bayes Estimates** en lugar de asumir su normalidad.

3.3 Modelos Lineales Generalizados (GLM)

Los Modelos Lineales Generalizados son una generalización de los modelos lineales para una variable respuesta perteneciente a la familia exponencial, en la que tenemos una función de enlace que describe como la media de la variable respuesta y la combinación lineal de variables explicativas están relacionadas. Según Paul Roback y Julie Regler en el libro *Beyond Multiple Linear Regression: Applied Generalized Linear Models and Multilevel Models in R* (Roback y Legler 2021), los GLM son una clase de modelos más amplia que tienen formas parecidas para sus varianzas, verosimilitudes y MLEs; generalizando la regresión lineal múltiple.

La **familia exponencial** suele tener esta forma:

$$f(y | \theta, \phi) = \exp \left[\frac{y\theta - b(\theta)}{a(\phi)} + c(y, \phi) \right]$$

En esta ecuación, θ es el **parámetro canónico** y representa la posición (location); mientras que ϕ es el **parámetro de dispersión** y representa la escala (scale). De la misma forma, a , b y c representan diferentes miembros de la familia exponencial. En función del parámetro de dispersión, podemos distinguir entre familias exponenciales de **un** parámetro, y familias exponenciales de **dos** parámetros.

Para determinar si un modelo está basado en un único parámetro θ , tenemos que poder escribir su función de probabilidad de la siguiente forma:

$$f(y; \theta) = e^{[a(y)b(\theta)+c(\theta)+d(y)]}$$

Si el conjunto de posibles valores de entrada no depende de θ , la familia exponencial será de un parámetro. Como familias exponenciales de un parámetro, tenemos las distribuciones de Poisson y la Binomial. Vamos a demostrar que la distribución de Poisson es, en efecto, una familia exponencial de un parámetro.

Para ello, aplicando propiedades logarítmicas, podemos definir la distribución de Poisson como:

$$P(Y = y) = e^{-\lambda} e^{y \log \lambda} e^{-\log(y!)} = e^{y \log \lambda - \lambda - \log(y!)}$$

Si comparamos esta función de masa de probabilidad con la función de probabilidad general para familias con un único parámetro, podemos ver que:

$$\begin{aligned}a(y) &= y \\b(\theta) &= \log(\lambda) \\c(\theta) &= -\lambda \\d(y) &= -\log(y!)\end{aligned}$$

La función $b(\theta)$ es lo que denominamos **enlace canónico**, una función que nos permite modelar como una función lineal de variables explicativas.

Como familias exponenciales de dos parámetros, tenemos la distribución Gamma y la Normal. De forma parecida a la anterior, podemos demostrar que la distribución Normal es una familia exponencial de dos parámetros.

Podemos definir la función de densidad de una distribución Normal como:

$$f(y|\mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(y-\mu)^2}{2\sigma^2}\right)$$

Si separamos términos y los escribimos como términos logarítmicos, tenemos que:

$$f(y|\mu, \sigma^2) = \exp\left(y \cdot \frac{\mu}{\sigma^2} - \frac{y^2}{2\sigma^2} + \left(-\frac{\mu^2}{2\sigma^2} - \frac{1}{2} \log(2\pi\sigma^2)\right)\right)$$

Si comparamos esta función de densidad con la forma general de la familia exponencial, podemos ver que:

$$\begin{aligned}a(y) &= y \\b(\mu, \sigma^2) &= \frac{\mu}{\sigma^2} \\c(\mu, \sigma^2) &= -\frac{\mu^2}{2\sigma^2} - \frac{1}{2} \log(2\pi\sigma^2) \\d(y, \sigma^2) &= -\frac{y^2}{2\sigma^2}\end{aligned}$$

Por lo tanto, demostramos que la distribución normal también pertenece a la familia exponencial, pero con una peculiaridad respecto a la distribución de Poisson: es una familia exponencial de dos parámetros, la media μ y la varianza σ^2 . En este caso, el término $b(\mu, \sigma^2)$ es el **enlace canónico** que conecta las variables explicativas con el modelo.

En concreto, para los casos en los que la respuesta no es normal, la ecuación del modelo es la siguiente:

$$g(E(y_{ij} | x_{ijk}, \beta_{0i}, \dots, \beta_{Ki})) = \beta_{0i} + \sum_{k=1}^K \beta_{ki} x_{ijk}$$

Donde g es la función enlace y, pese a que puede parecerse mucho a la función para modelos LMM, tienen algunas diferencias, como que en el primer miembro tenemos el enlace del valor

esperado en vez de la variable respuesta, y en el segundo miembro no se cuenta con los errores; por lo que no existe una matriz de correlaciones de los residuos. De esta forma, ya hemos **generalizado** nuestro modelo para manejar variables respuesta que no siguen una distribución normal. A través de esta generalización, somos capaces de escribir la función de masa o densidad de probabilidad de distintas distribuciones para poder modelar el enlace canónico como función lineal de las variables predictoras.

Una vez formulado el modelo GLM, los parámetros se estiman mediante el método de Máxima Verosimilitud (MLE). Para obtener los estimadores, se utiliza un procedimiento iterativo denominado Iteratively Reweighted Least Squares (IRLS), que ajusta los pesos de las observaciones en función de la varianza esperada.

En cuanto a la validación, se analizan medidas como la devianza (comparable a la suma de cuadrados en modelos lineales), el AIC para comparar modelos, y los residuos devianza y Pearson, que permiten detectar valores atípicos o mal ajuste. También es importante evaluar la distribución de la variable respuesta y el ajuste del modelo a través de gráficos de residuos.

3.3.1 Ejemplo práctico

Supongamos que queremos modelar el número de llamadas que recibe un centro de emergencias por hora (llamadas), en función del número de operadores de guardia (operadores). Este tipo de datos es típico de una distribución Poisson.

Call:

```
glm(formula = llamadas ~ operadores, family = poisson, data = df_poisson)
```

Coefficients:

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	0.65851	0.12904	5.103	3.34e-07 ***
operadores	0.25792	0.03495	7.380	1.58e-13 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for poisson family taken to be 1)

Null deviance: 138.738 on 99 degrees of freedom
 Residual deviance: 82.559 on 98 degrees of freedom
 AIC: 405.63

Number of Fisher Scoring iterations: 4

En este modelo, estamos modelando el logaritmo del número esperado de llamadas como una función lineal del número de operadores. La interpretación de los coeficientes es en términos del logaritmo de la tasa: por cada operador adicional, la tasa de llamadas aumenta multiplicativamente por $\exp(\cdot)$.

3.4 Modelos Lineales Generalizados Mixtos (GLMM)

Cuando trabajamos con datos longitudinales cuya variable respuesta no sigue una distribución normal, los Modelos Lineales Mixtos (LMM) dejan de ser apropiados. En estos casos, extendemos los modelos hacia los Modelos Lineales Generalizados Mixtos (GLMM), los cuales combinan la flexibilidad de los GLM con la estructura de efectos aleatorios de los LMM. Un GLMM permite modelar: variables respuesta que pertenecen a la familia exponencial (binomial, Poisson, Gamma...), y la correlación entre observaciones repetidas para el mismo individuo, mediante efectos aleatorios.

La ecuación general de un GLMM es:

$$g(E(y_{ij} | b_i)) = \beta_0 + \sum_{k=1}^K \beta_k x_{ijk} + Z_{ij} b_i$$

Donde:

- y_{ij} es la respuesta del individuo i en la ocasión j .
- x_{ijk} es el valor de la variable explicativa k para ese individuo en ese momento.
- β_{0i} es el intercepto aleatorio.
- β_{ki} puede incluir efectos fijos o aleatorios.
- $g(\cdot)$ es la función de enlace, que conecta el valor esperado con la combinación lineal de predictores.

Dependiendo de la naturaleza de y_{ij} , usaremos distintos enlaces. Para datos binarios: enlace logit y distribución binomial; para datos de conteo: enlace log y distribución de Poisson; y para tiempos o proporciones: enlaces adaptados como log-log, logit, etc. Si la variable respuesta es binaria (por ejemplo, éxito/fracaso), se usa un **modelo logístico mixto**:

$$\text{logit}(p_{ij}) = \beta_0 + \beta_1 x_{ij} + b_i$$

Donde:

- $\text{logit}(p) = \log\left(\frac{p}{1-p}\right)$
- $p_{ij} = \mathbb{P}(y_{ij} = 1 | b_i)$

- $b_i \sim \mathcal{N}(0, \sigma_b^2)$ es un efecto aleatorio por sujeto.

Esto permite modelar probabilidades condicionales considerando la variabilidad entre individuos.

3.4.1 Ejemplo práctico

Supongamos que queremos modelar si un estudiante aprueba un examen (`aprobado = 0/1`) en función de las horas de estudio (`horas`) y si el estudiante forma parte de un grupo diferente (`grupo`). Como los estudiantes tienen múltiples exámenes, añadimos un efecto aleatorio por estudiante.

```
Generalized linear mixed model fit by maximum likelihood (Laplace
Approximation) [glmerMod]
Family: binomial ( logit )
Formula: aprobado ~ horas + grupo + (1 | estudiante)
Data: df_glmm
```

AIC	BIC	logLik	deviance	df.resid
529.0	545.8	-260.5	521.0	496

Scaled residuals:

Min	1Q	Median	3Q	Max
-3.6296	-0.6712	0.3084	0.5411	2.1805

Random effects:

Groups	Name	Variance	Std.Dev.
estudiante	(Intercept)	0.8974	0.9473

Number of obs: 500, groups: estudiante, 100

Fixed effects:

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	-0.93155	0.27963	-3.331	0.000864 ***
horas	0.39577	0.05069	7.808	5.8e-15 ***
grupoB	0.26530	0.24055	1.103	0.270081

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Correlation of Fixed Effects:

	(Intr) horas
horas	-0.710
grupoB	-0.479 0.066

En este modelo GLMM (glmer) estamos modelando la probabilidad de aprobar un examen en función de dos variables explicativas: horas (número de horas de estudio) y grupo (grupo educativo (A o B), con A como categoría de referencia). Además, incluimos un efecto aleatorio de intercepto por estudiante, lo cual es adecuado porque cada estudiante tiene múltiples observaciones (exámenes), y esperamos que haya variabilidad entre ellos.

El intercepto aleatorio por estudiante tiene una desviación estándar de 0.9473, lo que indica que hay una variación importante en la propensión a aprobar entre estudiantes, incluso tras controlar por las horas de estudio y el grupo; lo que justifica el uso de un GLMM en lugar de un GLM clásico. Este modelo demuestra cómo un GLMM puede capturar variabilidad individual (entre estudiantes) y a la vez evaluar el efecto de factores fijos. El uso del modelo mixto es crucial: si ignoramos el efecto aleatorio por estudiante, estaríamos asumiendo que todos los estudiantes tienen la misma propensión a aprobar, lo cual claramente no es el caso según la varianza estimada.

Los GLMM tienen múltiples ventajas respecto a otros modelos, ya que permiten ajustar modelos a variables respuesta no continuas, incorporan variabilidad entre individuos mediante efectos aleatorios, se adaptan bien a datos longitudinales y jerárquicos, y permiten hacer inferencia poblacional y considerar la dependencia temporal. Por tanto, los GLMM constituyen una herramienta esencial para el análisis de datos longitudinales cuando la respuesta no es normal, ya que preservan la estructura de dependencia de los datos sin violar los supuestos del modelo.

La estimación de los parámetros en modelos GLMM es más compleja que en GLM o LMM, debido a que la función de verosimilitud no tiene una forma analítica cerrada. Por ello, se utilizan técnicas de aproximación numérica como: aproximación de Laplace, que integra los efectos aleatorios y aproxima la verosimilitud, cuadratura Gauss-Hermite adaptativa, que mejora la precisión en presencia de muchos efectos aleatorios, o Penalized Quasi-Likelihood (PQL), en versiones simplificadas. Estas aproximaciones permiten realizar inferencia sobre los efectos fijos y estimar la variabilidad entre sujetos.

Además, los GLMM pueden presentar problemas de convergencia, especialmente con estructuras complejas o tamaños muestrales pequeños. Es importante validar el modelo revisando los residuos, la bondad de ajuste (AIC, BIC), y la significación de los efectos aleatorios, por ejemplo, mediante test de razón de verosimilitudes anidados o comparaciones de modelos.

4 Análisis de la base de datos

El conjunto de datos World Happiness 2015-2024 recopila información sobre la felicidad percibida en diferentes países a lo largo de los años. Este dataset proviene de los informes anuales de felicidad publicados por la Red de Soluciones para el Desarrollo Sostenible de la ONU, los cuales se basan en encuestas realizadas a nivel mundial. El dataset tiene una buena cobertura temporal, ya que abarca datos de 2015 a 2024 y permite analizar tendencias a lo largo del tiempo, pero también geográfica, dado que incluye información de diferentes países y regiones del mundo. Este dataset es ampliamente utilizado en estudios de bienestar, calidad de vida y políticas públicas, y contiene métricas económicas y sociales que permiten un análisis estadístico y comparativo. Cada fila del dataset representa un país en un año determinado y contiene variables socioeconómicas y de bienestar que pueden influir en la percepción de felicidad de su población. Estas variables son:

- Country: Nombre del país.
- Region: Continente o agrupación geográfica del país.
- Happiness Score: Puntuación de felicidad promedio en el país (escala de 0 a 10).
- GDP per capita: Producto Interno Bruto per cápita ajustado por poder adquisitivo.
- Social Support: Medida de apoyo social basado en la percepción de las personas sobre la ayuda que pueden recibir de familiares y amigos.
- Healthy Life Expectancy: Esperanza de vida saludable en años.
- Freedom to Make Life Choices: Libertad para tomar decisiones personales, según encuestas de percepción.
- Generosity: Nivel de generosidad en la sociedad, basado en donaciones y ayuda a otros.
- Perceptions of Corruption: Nivel de percepción de corrupción en el gobierno y los negocios.

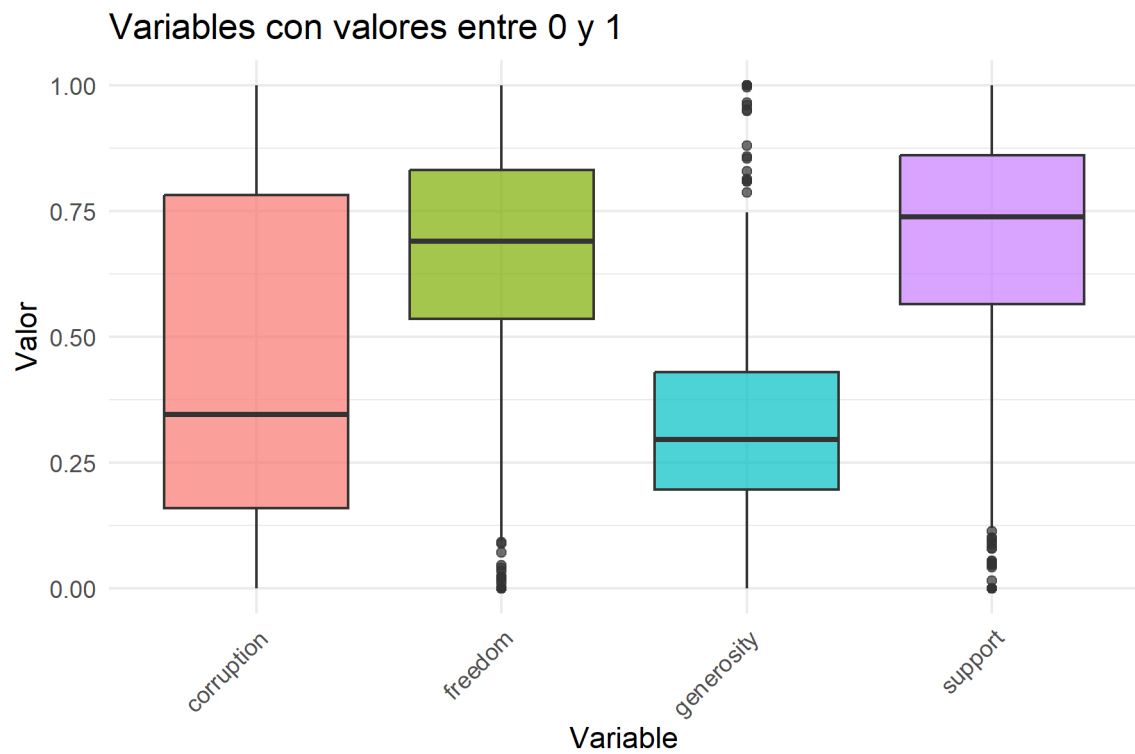


Figura 4.1: Distribución de las variables de la base de datos World Happiness

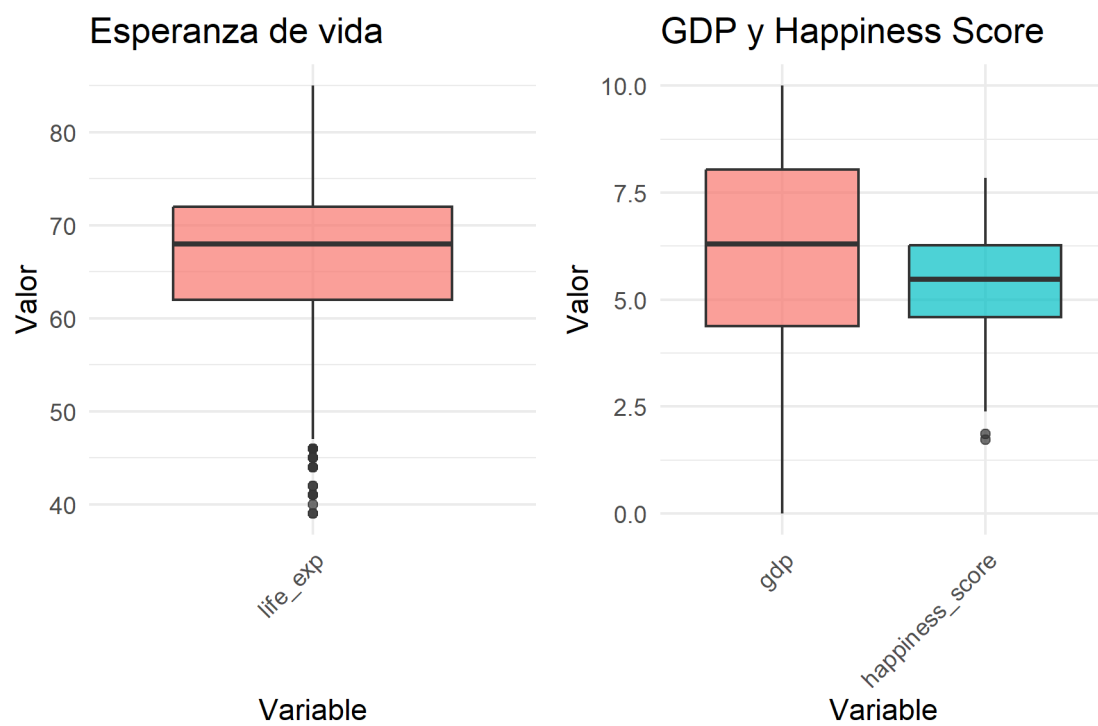


Figura 4.2: Distribución de las variables de la base de datos World Happiness

4.1 Análisis exploratorio inicial

Hemos comenzado el análisis de la base de datos, la cual no cuenta con datos faltantes. De la misma manera, estos gráficos de la Figura 4.1 y Figura 4.2 nos muestran la distribución de las diferentes variables, en la que se aprecian que tampoco hay valores atípicos. Explorando opciones para enriquecer la base de datos y, potencialmente, poder construir un modelo más informativo, se han estudiado diversas opciones y se ha decidido integrar información de dos fuentes externas que aportan indicadores políticos y de libertades civiles en los países. Estas bases de datos nos permitirán explorar hasta qué punto la democracia, los derechos políticos y las libertades influyen en la percepción de felicidad de las sociedades.

La primera base de datos que hemos considerado es “Freedom in the World” (Freedom House 2024), un informe anual de la organización Freedom House, que evalúa el estado de las libertades políticas y civiles en el mundo. Cada país es clasificado en función de indicadores de democracia, libertades individuales y derechos políticos. El motivo por el que hemos elegido esta base de datos es porque los estudios en ciencias sociales han mostrado que la percepción de felicidad no solo está ligada a factores económicos, sino también a la capacidad de los ciudadanos para expresarse libremente, participar en la política y vivir sin restricciones autoritarias. Incorporar estos datos nos permitirá ver si existe una correlación entre los niveles de libertad y la felicidad percibida en cada país. Como esta base de datos cuenta con una gran cantidad de variables, hemos elegido las siguientes variables de interés:

- Country/Territory: Identificación del país o territorio.
- Region: Indica la zona geográfica del país, similar al `regional_indicator` del dataset original.
- c/T: Diferencia entre países y territorios, aunque este concepto puede ser delicado según el análisis.
- Edition: Año del reporte, fundamental para el análisis longitudinal.
- Status: Clasificación del país en cuanto a su libertad: Libre (F), Parcialmente Libre (PF) o No Libre (NF).
- PR rating (Political Rights): Puntuación de 1 a 7 sobre derechos políticos.
- CL rating (Civil Liberties): Puntuación de 1 a 7 sobre libertades civiles.

La otra base de datos que hemos elegido para nuestro análisis es “Democracy Data”, un dataset que proviene del proyecto TidyTuesday y está basado en estudios académicos sobre democracia y regímenes políticos. Clasifica los países según su sistema de gobierno y proporciona información detallada sobre su historia política. Una de las características que tiene esta base de datos es que tiene datos hasta 2020, por lo que tenemos que considerar que, si vamos a trabajar con ella, tendremos las características completas de las observaciones en un período reducido. Dado que la felicidad no solo depende de factores económicos, sino también de la estabilidad política y la gobernanza, estas variables pueden ayudarnos a explicar por qué algunos países presentan niveles bajos de felicidad a pesar de tener una economía sólida. Al igual que en el caso anterior, como esta base de datos contiene más de 40 variables,

hemos decidido quedarnos con aquellas que pensamos que mejor se adaptan a nuestro análisis. Estas variables son:

- `country_name`: Nombre del país.
- `year`: Año de observación.
- `regime_category`: Clasificación del sistema de gobierno (democracia parlamentaria, autocracia civil, dictadura militar, monarquía, etc.).
- `is_monarchy`: Indica si el país es una monarquía.
- `is_democracy`: Indica si el país es una democracia.
- `is_presidential`: Indica si el sistema es presidencialista.
- `is_colony`: Identifica si el país sigue siendo una colonia.
- `is_communist`: Indica si el país sigue un régimen comunista.
- `spatial_democracy`: Evalúa el nivel de democracia en los países vecinos.
- `has_full_suffrage`: Indica si hay sufragio universal.
- `electoral_category`: Tipo de elecciones (no democráticas, de partido único, multipartidistas no democráticas o democráticas).
- `spatial_electoral`: Evalúa la calidad electoral de los países vecinos.
- `has_free_and_fair_election`: Indica si las elecciones son libres y justas.
- `has_alternation`: Indica si existe alternancia en el poder.

Agregando estos datasets consideramos que se puede abarcar de forma bastante completa los diferentes factores que afectan a la percepción de la felicidad. Esto nos permitirá realizar un análisis más profundo y responder preguntas sobre la relación entre política, democracia y bienestar subjetivo en distintos países.

4.2 Análisis de las bases de datos complementarias

Al trabajar con datos longitudinales a nivel país, surge una cuestión clave: ¿en qué medida las características políticas de cada país se mantienen estables o experimentan cambios relevantes a lo largo del tiempo? En este estudio, incorporamos información de las bases de datos Freedom in the World y Democracy Data, que permiten capturar distintos aspectos del contexto político de los países, como el tipo de régimen, la alternancia en el poder o la existencia de elecciones libres.

Sin embargo, nos encontramos con dos limitaciones importantes. En primer lugar, muchas de estas variables presentan una cobertura temporal limitada, con registros ausentes en varios años dentro del periodo 2015–2023. En segundo lugar, se trata de variables que tienden a ser bastante estables: en la mayoría de los países analizados no se observan cambios relevantes a lo largo del tiempo. Estas circunstancias dificultan su tratamiento como verdaderas variables longitudinales.

Por ello, optamos por incorporar estas variables como una “foto fija” del contexto político, extrayendo su valor más representativo del periodo estudiado. De esta forma, las utilizamos

como covariables fijas en los modelos de felicidad, contribuyendo a enriquecer el análisis sin introducir sesgos derivados de valores faltantes o variaciones mínimas que podrían ser ruido.

No obstante, existen excepciones interesantes. Algunos países sí han experimentado cambios significativos en su sistema político, su forma de gobierno o su nivel de democracia. En la siguiente sección se representa gráficamente la evolución de dichas variables para estos casos específicos, con el objetivo de ilustrar de forma visual estos procesos de cambio.

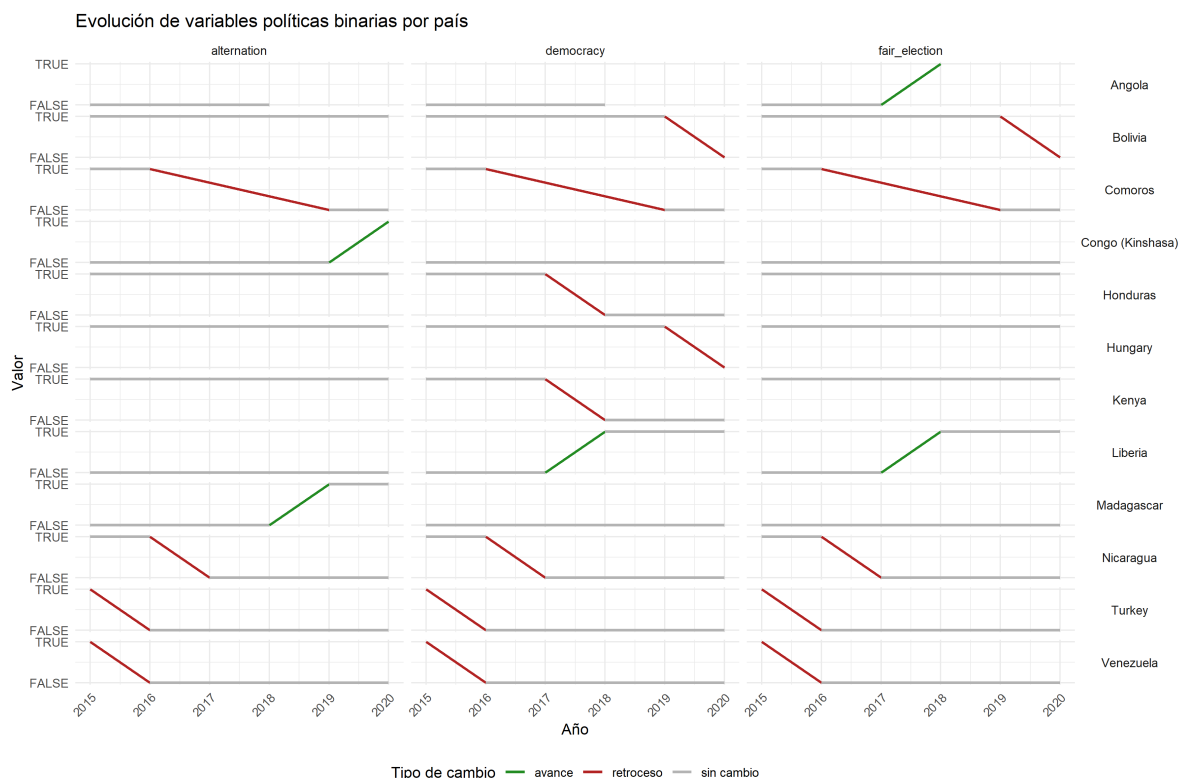


Figura 4.3: Evolución temporal de las variables políticas binarias de la base de datos Democracy Data

Estos gráficos sobre la evolución de diferentes variables políticas muestran de forma clara y visual la evolución temporal de varios países que han experimentado cambios entre 2015 y 2020.

Si observamos los cambios en el tipo de elecciones en la Figura 4.5, podemos ver que Venezuela presentó hasta 3 categorías distintas: pasó de elecciones democráticas a elecciones no democráticas multi-partido y luego a no tener elecciones directamente, reflejando el colapso progresivo del sistema electoral bajo el régimen de Maduro. Otros países como Honduras o Comoras alternan entre elecciones democráticas y elecciones no democráticas multi-partido, reflejando cuestionamientos sobre la transparencia y legitimidad de sus procesos electorales

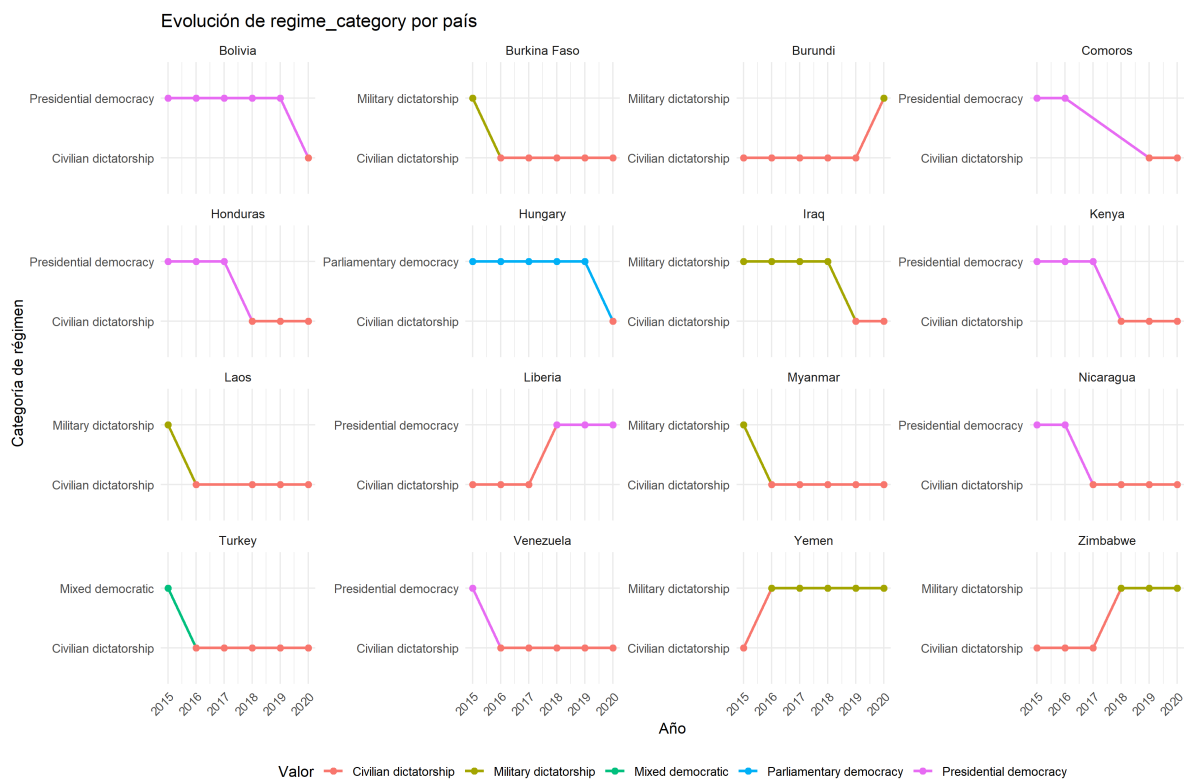


Figura 4.4: Evolución del cambio de régimen democrático en países desde 2015 a 2020

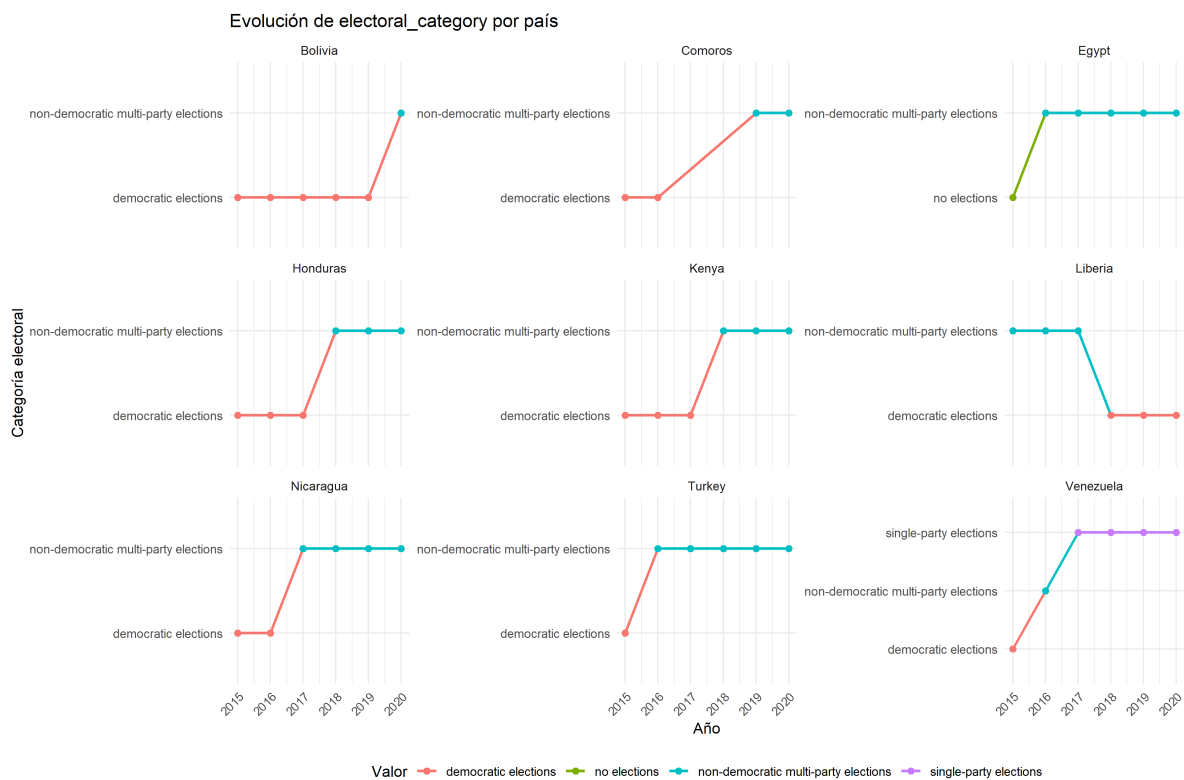


Figura 4.5: Evolución del cambio de categoría electoral en países desde 2015 a 2020

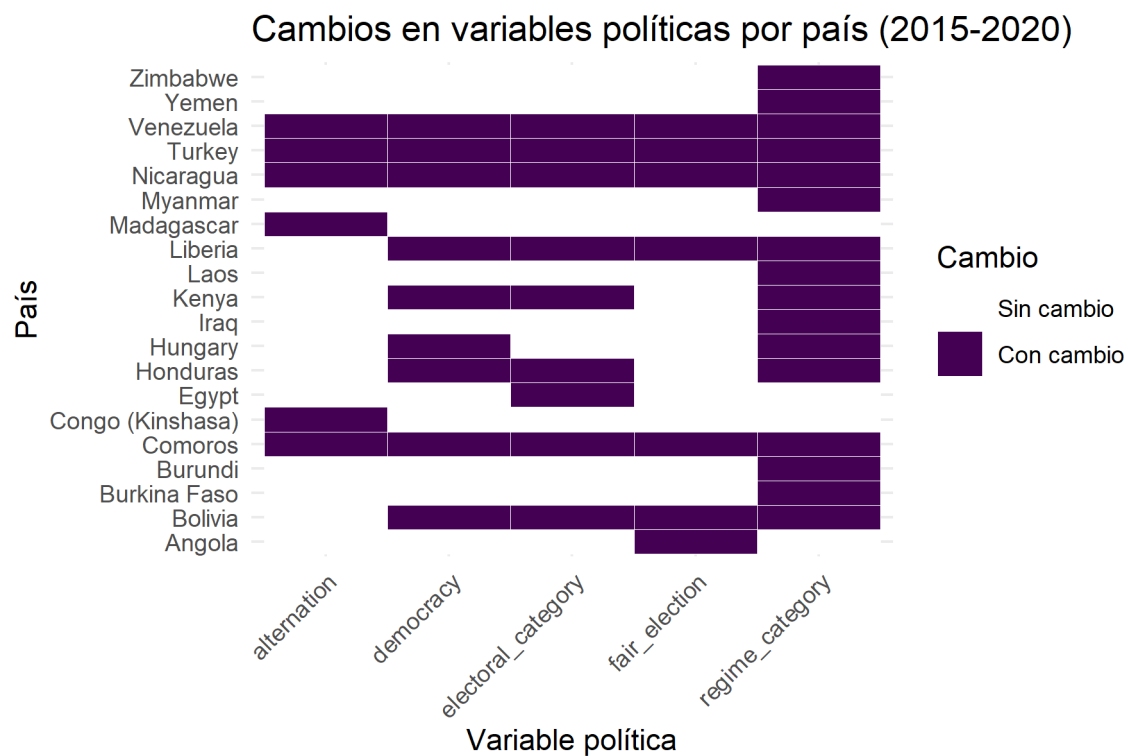


Figura 4.6: Análisis de los distintos cambios políticos en países de 2015 a 2020

y evidenciando las dificultades para establecer un sistema electoral pluralista. En el caso de Honduras a partir de 2018, lo que coincide con la reelección altamente cuestionada de Juan Orlando Hernández, rodeada de denuncias de fraude y represión; mientras que Comoras en 2019, tras una serie de reformas constitucionales que favorecieron la consolidación del poder del presidente Azali Assoumani.

En cuanto a los cambios en la alternancia en el poder, los cuales podemos apreciar en la Figura 4.3, vemos cómo en Nicaragua y Venezuela esta variable refleja la falta de alternancia, mostrando un patrón de concentración del poder ya que en años anteriores había alternancia, pero se pierde conforme los gobiernos se consolidan en el poder.

Siguiendo este mismo gráfico, podemos observar los cambios en la existencia de elecciones libres y justas, en los que vemos que Angola y Liberia reflejan mejoras en sus procesos electorales con la celebración de elecciones que fueron más competitivas y abiertas que en años anteriores; pero hay otros casos como Bolivia, que presenta un cambio negativo en 2019, año marcado por acusaciones de fraude electoral que desembocaron en la crisis política. Otro ejemplo que se puede destacar es Turquía, que deja de tener elecciones libres y justas en los últimos años, en línea con el creciente autoritarismo del presidente Erdoğan.

Para acabar con el Figura 4.3, observando la condición de democracia, vemos que Nicaragua dejó de ser considerada una democracia a partir de 2017, cuando se intensificó el control autoritario de Daniel Ortega sobre las instituciones y se reprimieron masivamente las protestas. Hungría también dejó de ser considerada democracia a partir de 2019, en consonancia con el cambio de régimen antes mencionado. Otro cambio lo observamos en Bolivia, que fue considerada democracia durante gran parte del período, pero dejó de serlo temporalmente tras la crisis política y la salida de Evo Morales en 2019.

Si observamos la forma de gobierno presidencial en la Figura 4.6, vemos que el único Estado que cambia es China: aunque siempre fue una dictadura de partido único, el dataset refleja un cambio en esta variable posiblemente por una reinterpretación metodológica, ya que a partir de 2019 es considerada presidencial. Es destacable que esta variable apenas presenta cambios (salvo China), lo cual sugiere que la forma de gobierno en términos institucionales es más estable que otras dimensiones como la calidad electoral o la democracia sustantiva. Esto puede ayudar a entender que muchos procesos de regresión democrática no implican reformas estructurales del sistema político, sino más bien una transformación interna de las reglas del juego bajo el mismo marco institucional.

Por último, si observamos los cambios en el tipo de régimen en la Figura 4.4, vemos como Hungría pasó de ser una democracia parlamentaria a una dictadura civil, lo cual refleja el creciente autoritarismo del gobierno de Viktor Orbán y el deterioro de la separación de poderes y la libertad de prensa. Sin embargo, hay casos como Liberia que sucede todo lo contrario: en los primeros años del periodo considerado, Liberia fue clasificada como una dictadura civil, reflejo del mandato de Ellen Johnson Sirleaf, que, aunque existía cierta estructura institucional, no se garantizaban plenamente principios básicos de la democracia como elecciones libres o alternancia real de poder. Sin embargo, a partir de 2018, Liberia pasa a ser clasificada como una

democracia presidencial, cambio que coincide con la llegada al poder del exfutbolista George Weah, quien fue elegido democráticamente en un proceso que representó la primera transición pacífica de poder entre dos presidentes electos en el país desde 1944. Esta transición marcó un punto de inflexión político, reflejando una mejora significativa en las instituciones democráticas del país. En Burundi, el cambio en el tipo de régimen refleja el deterioro institucional que siguió a la decisión de Pierre Nkurunziza de presentarse a un tercer mandato en 2015, desatando una grave crisis política. El país pasa a ser clasificado como dictadura civil, reflejando la suspensión de libertades y el cierre del espacio cívico. Este cambio es particularmente relevante porque anticipa un patrón que luego se repite en otros países: el uso del poder electoral como instrumento de legitimación de autoritarismo. Otro caso a destacar es Zimbabwe, ya que es interesante porque, aunque solo presenta un cambio en `regime_category`, este se produce tras la caída de Robert Mugabe en 2017, una transición que fue vista por algunos sectores como una oportunidad para abrir un nuevo ciclo democrático. Sin embargo, el hecho de que no haya cambios en otras variables como `is_democracy` o `has_free_and_fair_election` refleja que el cambio de liderazgo no implicó necesariamente una mejora sustantiva en la calidad del régimen.

Además de los ejemplos más visibles, hay otros países que presentan cambios más sutiles pero igualmente significativos. Por ejemplo, Madagascar muestra una alteración en la variable `has_alternation`, indicando un momento de alternancia que podría asociarse con las elecciones de 2018, cuando Andry Rajoelina volvió al poder tras haberlo ocupado anteriormente como presidente de transición. Este cambio sugiere un contexto político volátil en el que las alternancias no siempre reflejan procesos plenamente democráticos, sino que pueden estar vinculadas a luchas internas o acuerdos de élites.

Otro caso particular es Comoras, que presenta múltiples cambios en variables como el tipo de elecciones, la democracia y el tipo de régimen. Esto refleja un proceso de consolidación autoritaria que ha sido documentado por organismos internacionales tras la reforma constitucional de 2018 y las elecciones de 2019, en las que se concentró el poder presidencial y se limitó la oposición. Comoras es un ejemplo claro de regresión democrática multidimensional, donde no solo se pierde calidad electoral, sino que se transforman también las estructuras institucionales.

En Kenya, se observan cambios tanto en `is_democracy` como en `electoral_category` y `regime_category`, lo cual es coherente con un contexto de avances y retrocesos en la calidad democrática. A pesar de tener elecciones relativamente regulares, las denuncias de fraude, violencia post-electoral y polarización han afectado la credibilidad del sistema. El dataset parece capturar bien esa inestabilidad institucional, mostrando a Kenya como un país en constante disputa entre aperturas democráticas y tendencias autoritarias.

Este gráfico refuerza la idea de que varios países han experimentado retrocesos democráticos significativos, especialmente en torno a elecciones libres, alternancia y la clasificación del régimen. Además, evidencia que estos cambios no son simultáneos: mientras algunos países cambian en 2016, otros lo hacen en 2018 o 2020, lo cual permite contextualizar los cambios políticos con eventos históricos concretos en cada nación. Este tipo de análisis temporal no

solo es útil para identificar patrones políticos, sino también para cruzar estos cambios con la percepción de felicidad de la población y evaluar si existe alguna asociación relevante entre ambas dimensiones.

Dado que uno de los objetivos principales de este análisis es estudiar la evolución de la felicidad y sus determinantes a lo largo del tiempo, necesitamos trabajar con una base de datos que tenga cobertura completa para el periodo 2015–2024. En este sentido, hemos optado por utilizar como base principal la combinación de los datasets World Happiness y Freedom in the World, ya que ofrecen una continuidad temporal adecuada y permiten aplicar técnicas de medidas repetidas para estudiar cómo varían las observaciones de un mismo país a lo largo de varios años.

Por su parte, la base de datos Democracy Data presenta una limitación temporal importante, al contener información solo hasta el año 2020. Además, muchas de sus variables muestran escasa variación a lo largo del tiempo, lo que refuerza la idea de que no se comportan como verdaderas variables longitudinales.

Por ello, en lugar de excluir completamente esta base, la integramos como complemento, utilizándola para construir una fotografía del contexto institucional y político de cada país. Esta fotografía nos resulta especialmente útil a la hora de interpretar fenómenos detectados en el análisis longitudinal, como caídas o aumentos abruptos en el nivel de felicidad. En esos casos, el tipo de régimen, la alternancia en el poder o la existencia de elecciones libres pueden aportar claves importantes para entender estos cambios.

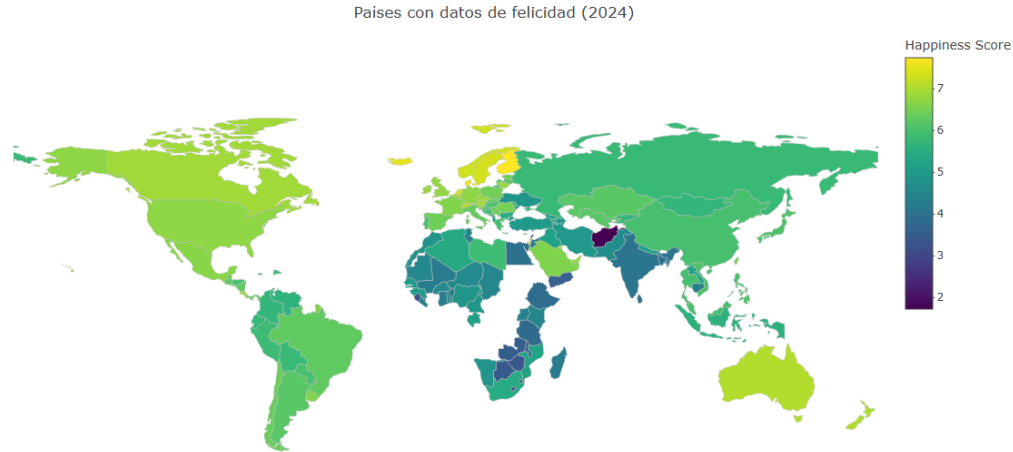
Antes de comenzar con la visualización de la evolución del índice de felicidad, resulta pertinente detenernos brevemente en la base Freedom in the World, que hemos decidido conservar por su cobertura temporal completa y por ofrecer indicadores relevantes sobre libertades civiles y derechos políticos. Aunque sus variables presentan cierta estabilidad en muchos países, hemos identificado diferencias notables en algunos casos que pueden ser útiles para contextualizar cambios observados en la felicidad. Además, no se detectan valores atípicos preocupantes ni problemas graves de calidad en la información, por lo que resulta razonable integrarla en el análisis longitudinal.

Una vez preparado el dataset final, es posible explorar cómo ha evolucionado la variable objetivo —`happiness_score`— entre 2015 y 2024. Para ello, se recurre a herramientas visuales como mapas o gráficos de línea por región y país, que permiten detectar tanto patrones generales como fluctuaciones específicas en determinados territorios. Estas visualizaciones constituyen el primer acercamiento al análisis longitudinal y ofrecen pistas sobre posibles factores explicativos de los cambios observados.

4.3 Evolución de la felicidad a lo largo del tiempo

Con el objetivo de identificar patrones geográficos en los niveles de felicidad, se emplean mapas coropléticos que permiten visualizar espacialmente la distribución del `happiness_score` a lo

largo del tiempo. Este enfoque facilita detectar agrupaciones regionales, contrastes entre países vecinos y posibles focos de evolución atípica. Al representar la información sobre el territorio, los mapas se convierten en una herramienta clave para complementar el análisis longitudinal y contextualizar las dinámicas regionales de la felicidad.



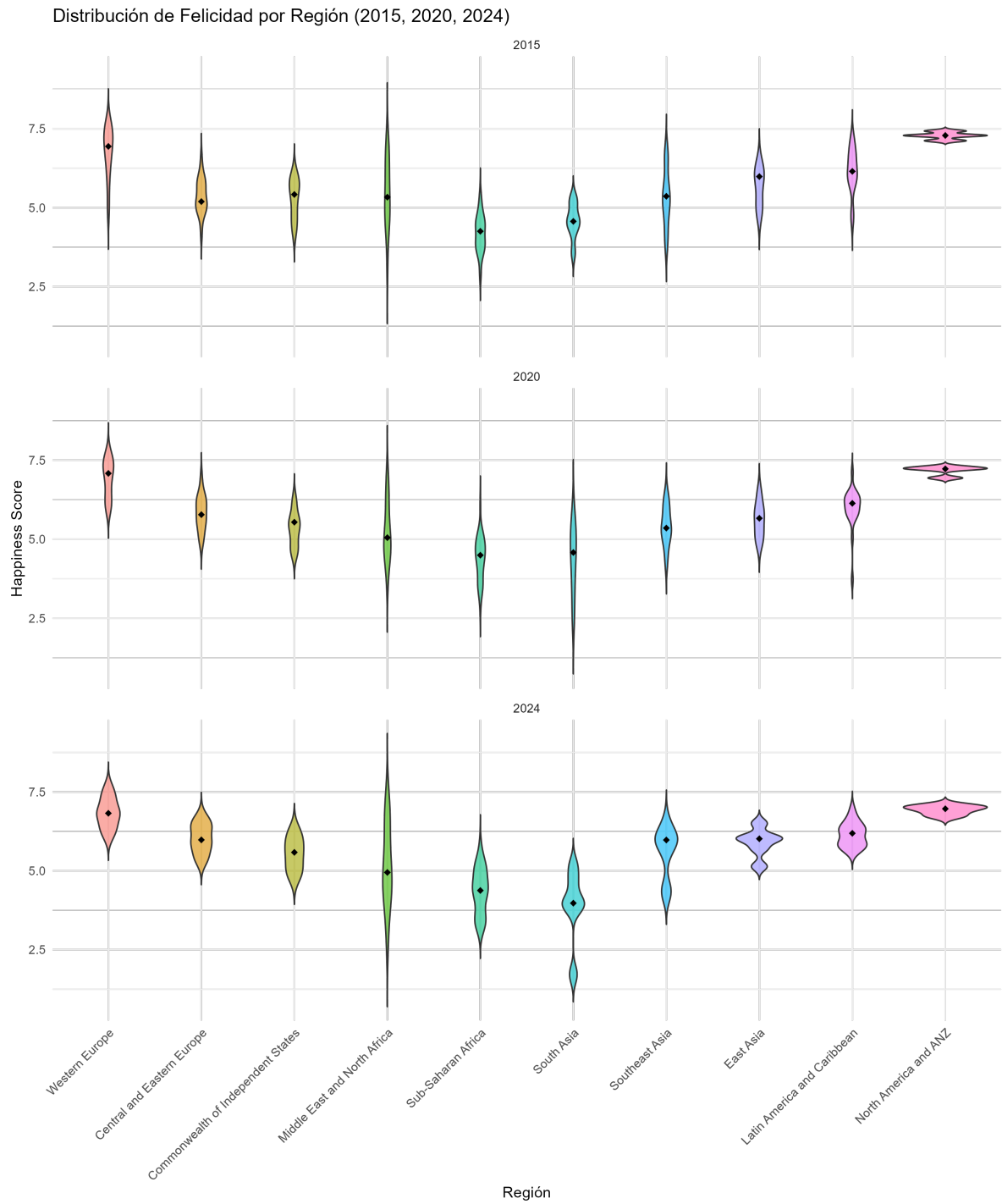
El mapa de felicidad por país en 2024 muestra una clara heterogeneidad geográfica en la percepción del bienestar, capturada a través del Happiness Score. Se observan regiones del norte global, como Europa Occidental, América del Norte y Oceanía, con los niveles más altos de felicidad (colores amarillos y verde claro), mientras que regiones del sur global, especialmente África Subsahariana y algunas partes de Asia como Afganistán o India, presentan valores notablemente más bajos (colores azul oscuro o morado).

Países como Finlandia, Noruega y Nueva Zelanda destacan con las puntuaciones más altas, reflejando contextos estables y altos niveles de desarrollo económico y social. En contraste, países como Zimbabwe, Sudán del Sur o Afganistán presentan los niveles más bajos, lo cual es coherente con contextos de conflicto, pobreza o inestabilidad política.

Este mapa no solo permite identificar diferencias entre regiones, sino también subrayar patrones estructurales: América Latina, por ejemplo, muestra un nivel medio de felicidad, con cierta variabilidad entre países. Es especialmente útil para detectar casos anómalos, como países con puntuaciones bajas en regiones generalmente altas o viceversa.

En conjunto, el gráfico evidencia de forma visual el impacto que pueden tener factores estructurales como el desarrollo, la gobernanza o la estabilidad en la percepción subjetiva del bienestar de los ciudadanos a nivel global.

Además del análisis temporal por país, resulta de interés observar cómo varía la felicidad entre regiones, y cómo se distribuye internamente en cada una de ellas. Para ello, se emplean gráficos de violín, que permiten visualizar simultáneamente la dispersión, la mediana y la densidad de los valores de `happiness_score` en cada región. Esta representación facilita la comparación entre regiones y la detección de desigualdades internas que podrían pasar desapercibidas en un análisis puramente temporal.



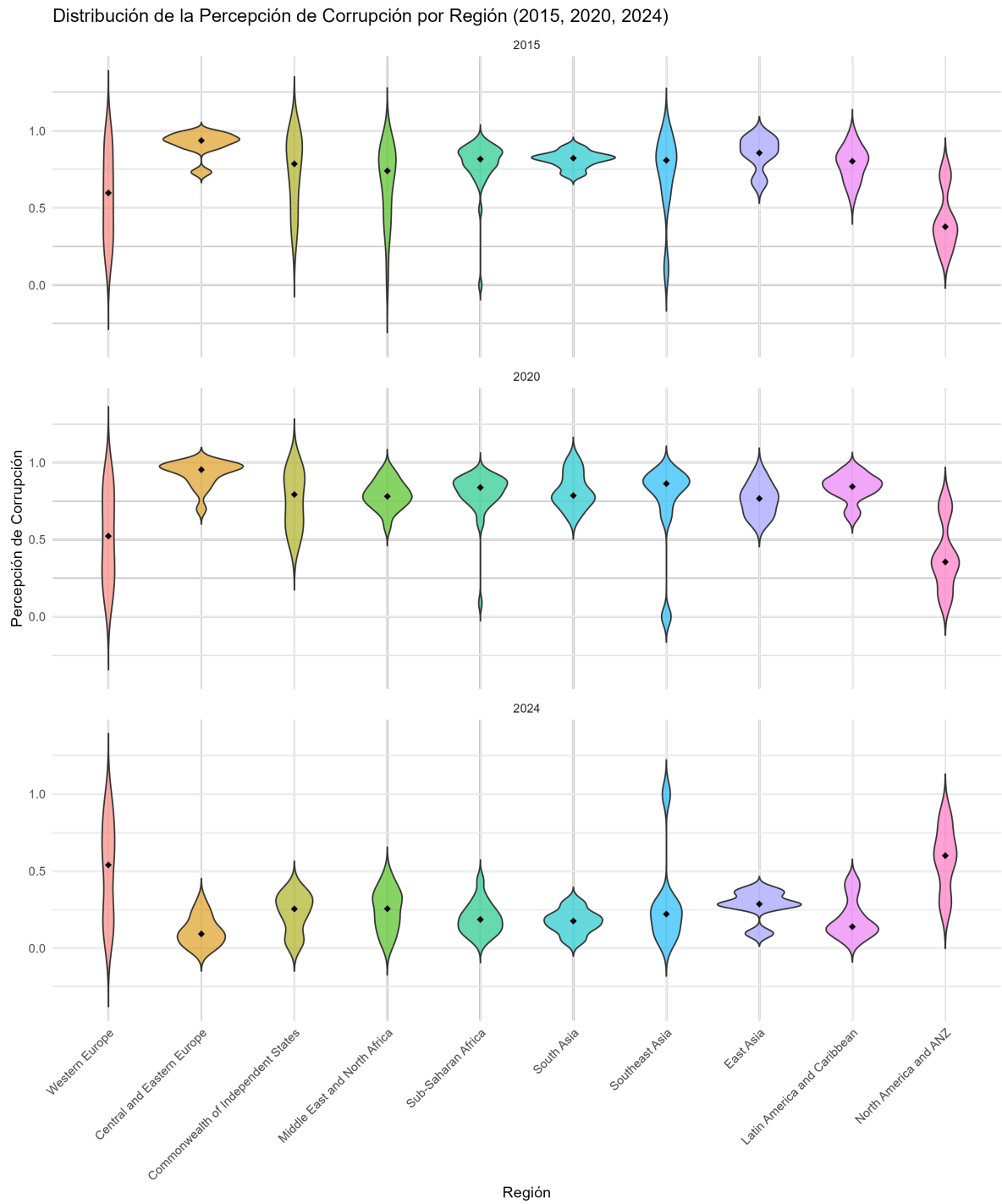
Observando la distribución de la felicidad por región, si nos fijamos en la mediana, que se muestra con un rombo negro en cada violín, vemos que hay regiones como Europa

Occidental o Norte América y Australia que mantienen medianas altas y estables a lo largo del tiempo, mientras que el Sur de Asia y África Sub-Sahariana tienen medianas bajas, aunque relativamente estables.

En estos gráficos también se puede observar cierta bimodalidad, ya que algunas regiones, como Latino América y el Caribe o el Oriente Medio y África del Norte muestran dos modos (zonas más anchas del violín), lo que sugiere heterogeneidad interna: algunos países con altos niveles de felicidad conviven con otros con puntuaciones mucho más bajas. Asia del Sur presenta una distribución muy desigual en 2024, con un notable ensanchamiento en la parte inferior del gráfico, reflejando una baja felicidad en ciertos países (como Afganistán).

En general, la forma de las distribuciones se mantiene similar entre 2015 y 2024 en muchas regiones, pero algunas muestran mayor dispersión, sobre todo en el continente asiático, ya que Oriente Medio y Asia del Sur sufren una caída considerable en la felicidad. Regiones como Europa Occidental y Norte América y Australia presentan distribuciones compactas y felicidad más alta y estable, lo que refleja mayor consistencia en el bienestar subjetivo de sus poblaciones.

A continuación, analizamos la percepción de la corrupción, una variable que diversos estudios han señalado como influyente en la percepción de felicidad. En la siguiente figura se representa la distribución de esta variable mediante gráficos de violines, lo que nos permite observar sus diferencias entre regiones. Este tipo de representación facilita comparar no solo los valores centrales, sino también la dispersión y asimetría en cada grupo geográfico. Se observa que algunas regiones presentan niveles percibidos de corrupción notablemente más bajos, lo que podría estar asociado a mayores niveles de bienestar subjetivo.



El gráfico de violines permite observar la evolución de la percepción de corrupción en las distintas regiones a lo largo del tiempo, mostrando tanto la mediana como la distribución

completa de los datos. Lo más destacable es la aparición de una clara bimodalidad en algunas regiones a partir de 2020 y especialmente en 2024, lo que sugiere una divergencia creciente entre países dentro de la misma región.

Europa Occidental mantiene una distribución relativamente estable a lo largo del tiempo, con una percepción de corrupción baja en general (valores altos), aunque en 2024 aparece una mayor dispersión que en 2015 y 2020. Sin embargo, en Europa Central y Oriental se observa un patrón inverso. En 2015 y 2020, la percepción es bastante alta (valores cercanos a 1), pero en 2024 se evidencia una caída brusca en la percepción positiva (valores muy bajos), lo que indica un empeoramiento de la percepción ciudadana sobre la corrupción en algunos países, como puede ser el caso de Hungría o Polonia.

Commonwealth of Independent States muestra una bimodalidad clara en 2024, mientras que en 2015 y 2020 predominaba un grupo más homogéneo. Esto sugiere que algunos países han experimentado mejoras o retrocesos significativos, diferenciándose del resto. América Latina y el Caribe también muestra un cambio claro: de una percepción moderadamente positiva en 2015 a una distribución más polarizada en 2020 y 2024, reflejando posiblemente tensiones políticas y escándalos de corrupción en países como Venezuela, Nicaragua o Bolivia.

En Asia del Este se observa una distribución muy compacta en 2015 y 2020, pero con una clara ruptura en 2024, lo que podría estar relacionado con dinámicas opuestas en países como China (control político férreo) frente a otros con gobiernos más abiertos. Por el contrario, regiones como Asia del Sur y África Sub-Sahariana muestran distribuciones bajas y más estables, aunque con una ligera mejora hacia 2024 en la percepción media.

Este análisis permite detectar no solo la evolución de la percepción media, sino también la heterogeneidad creciente dentro de regiones, y puede servir de base para cruzar estos patrones con los cambios políticos identificados previamente, evaluando si el deterioro institucional y la concentración del poder están efectivamente correlacionados con una mayor percepción de corrupción entre la ciudadanía.

Con el objetivo de comprender cómo ha cambiado la situación global en términos de bienestar subjetivo y factores asociados, analizamos a continuación la evolución promedio anual de las principales variables del informe de felicidad. Esta visualización nos permitirá detectar tendencias crecientes o decrecientes y posibles impactos globales, como crisis políticas, sanitarias o económicas.

La evolución anual del promedio de variables de interés de la Figura 4.7 refleja tendencias globales relevantes, pero su comparación con países concretos permite matizar dichas dinámicas y comprender mejor cómo contextos políticos específicos afectan al bienestar percibido por la población. Para ello, hemos seleccionado una serie de países que destacan por haber experimentado eventos políticos significativos o cambios institucionales durante el periodo de estudio, como Venezuela, Turquía o Hungría. Estos casos permiten ilustrar cómo la evolución de las condiciones políticas puede influir en la trayectoria del bienestar subjetivo, aportando una dimensión interpretativa más rica al análisis global.

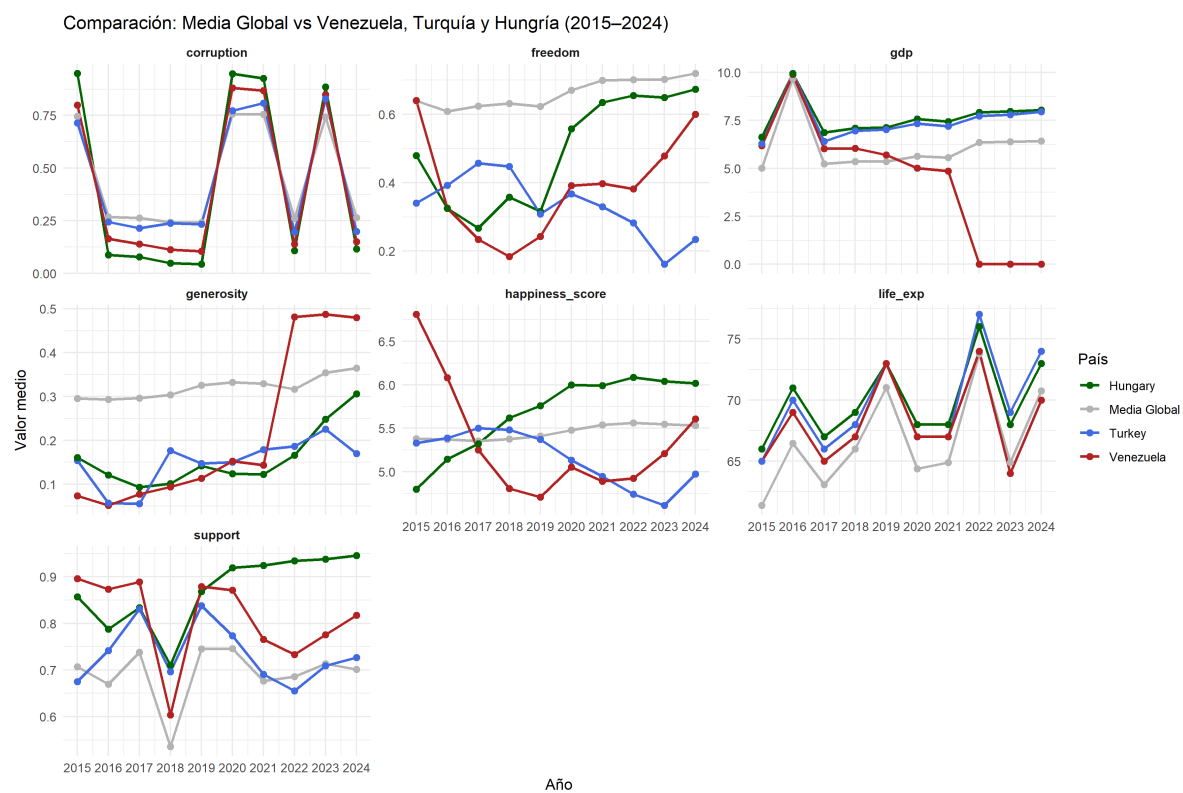


Figura 4.7: Análisis temporal de las principales variables del informe de felicidad en comparación con Hungría, Turquía y Venezuela

En primer lugar, Venezuela se desmarca de forma clara en varias variables. El PIB per cápita presenta una caída drástica desde 2017, situándose muy por debajo de la media global a partir de 2018, lo que refleja la aguda crisis económica que atraviesa el país. Esta caída va acompañada de un descenso en la libertad para tomar decisiones y en la felicidad percibida, lo que sugiere un deterioro generalizado del bienestar. Sin embargo, resulta llamativo el aumento pronunciado en la generosidad a partir de 2021, superando incluso a la media global. Este fenómeno podría estar vinculado a la solidaridad comunitaria surgida ante la crisis prolongada, así como a cambios metodológicos en la forma en la que se capta esta variable en contextos de alta inestabilidad.

Por otro lado, Turquía muestra una tendencia decreciente en la libertad para tomar decisiones, especialmente desde 2018, coincidiendo con el fortalecimiento del poder ejecutivo bajo el liderazgo de Recep Tayyip Erdoğan. A pesar de mantener un PIB per cápita relativamente estable y una esperanza de vida en línea con la media global, la percepción de corrupción se mantiene sistemáticamente alta, lo cual refuerza la idea de un deterioro institucional progresivo. La felicidad percibida en Turquía permanece por debajo de la media, lo que puede reflejar un desencanto social persistente.

En el caso de Hungría, se observa una situación más ambivalente. Por un lado, variables como la esperanza de vida, el apoyo social y la felicidad percibida se mantienen por encima de la media global, lo que sugiere una cierta estabilidad material y comunitaria. No obstante, la libertad para tomar decisiones experimenta un estancamiento, y la percepción de corrupción es elevada, lo cual coincide con el proceso de retroceso democrático documentado en el país desde mediados de la década de 2010. Este contraste pone de manifiesto cómo un país puede mantener ciertos niveles de bienestar mientras erosiona sus instituciones democráticas.

Este análisis complementa el estudio previo de la evolución global. Ya habíamos observado que, en conjunto, variables como la libertad para tomar decisiones y la generosidad presentan una evolución creciente, posiblemente impulsada por procesos de recuperación post-pandemia o cambios culturales. También destacaba un pico atípico en el PIB per cápita en 2016, probablemente debido a valores extremos, y una tendencia general de aumento en la felicidad percibida hasta 2022, seguida de un leve descenso. Por el contrario, la percepción de corrupción es más errática, con oscilaciones abruptas entre años, lo que refleja fuertes diferencias entre países.

La comparación entre países permite, por tanto, observar cómo procesos políticos específicos (autoritarismo, crisis económica, reformas institucionales) se reflejan en los indicadores de bienestar. Venezuela, Turquía y Hungría muestran trayectorias divergentes respecto a la media global, lo cual subraya la necesidad de un análisis desagregado para captar la complejidad del bienestar en contextos políticamente inestables o en transformación.

La Figura 4.8 muestra la evolución del Happiness Score por país entre los años 2015 y 2024, representando cada año en un panel independiente y añadiendo como referencia una línea roja discontinua que indica la media global de felicidad de ese año. Esta visualización permite

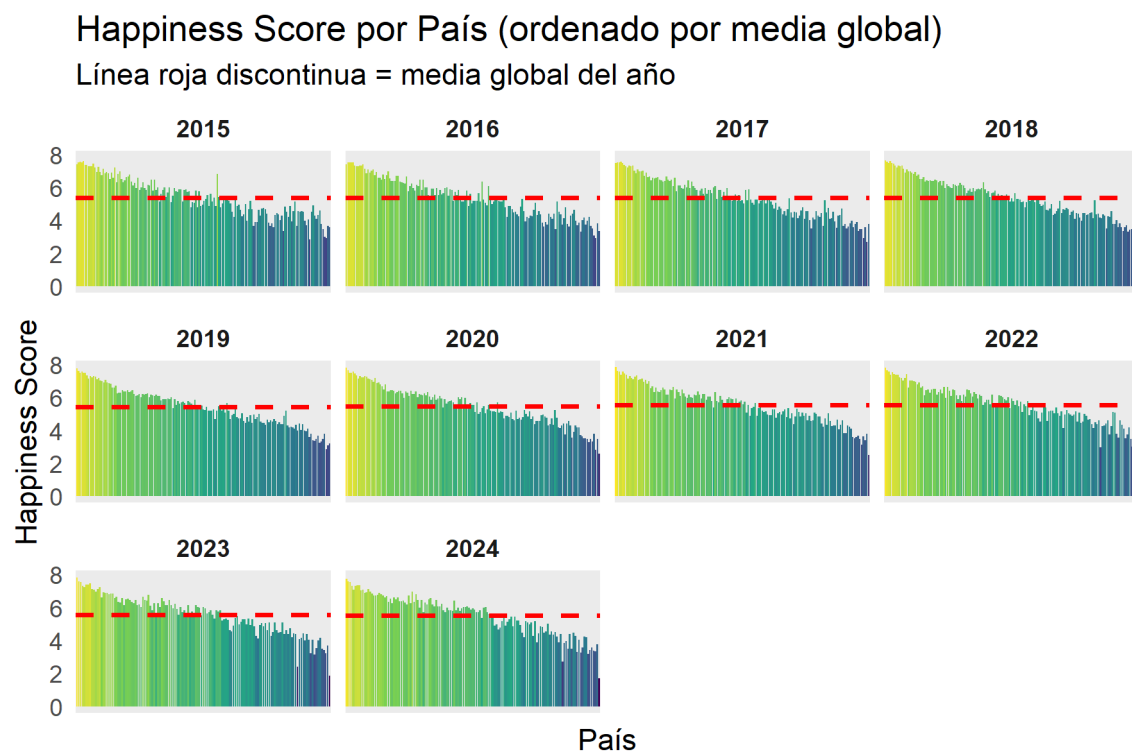


Figura 4.8: Análisis temporal de la evolución temporal de la felicidad por país

realizar un análisis comparativo tanto entre países como a lo largo del tiempo, aportando una perspectiva clara sobre la distribución del bienestar subjetivo a escala mundial.

Lo primero que llama la atención es la notable estabilidad de la media global a lo largo del periodo observado. La línea roja, que representa dicho promedio, se mantiene en un intervalo relativamente estrecho —aproximadamente entre 5.3 y 5.5 puntos—, lo que sugiere que, a nivel agregado, la percepción de felicidad mundial no ha sufrido cambios drásticos. Sin embargo, esta estabilidad global oculta dinámicas internas más complejas.

A medida que avanzan los años, especialmente a partir de 2020, se observa un aumento progresivo en la cantidad de países que se sitúan por debajo de la media global. Esta tendencia podría estar vinculada al impacto de la pandemia de COVID-19 y sus secuelas económicas, sanitarias y sociales, que afectaron de manera desigual a distintas regiones. Asimismo, podría reflejar procesos de deterioro democrático o inestabilidad política en determinados contextos, tal y como se ha explorado en análisis previos sobre variables institucionales.

El gráfico también pone de manifiesto la persistencia de fuertes desigualdades en la distribución del bienestar subjetivo. Mientras que un pequeño grupo de países —posiblemente naciones del norte de Europa como Finlandia, Dinamarca o Islandia— se mantiene de forma consistente en la parte superior del ranking con puntuaciones muy por encima de la media, la gran mayoría de países se concentra en la franja media o baja. Esta asimetría en la distribución revela que, si bien algunos países gozan de altos niveles de satisfacción vital sostenidos en el tiempo, muchos otros enfrentan barreras estructurales que dificultan mejorar su puntuación de felicidad.

Un fenómeno adicional que merece atención es el ligero descenso en los valores máximos del Happiness Score en los últimos años. En 2023 y 2024, incluso los países mejor posicionados parecen haber perdido algo de ventaja con respecto a años anteriores. Esta caída podría interpretarse como un efecto retardado de los acontecimientos globales recientes, o como una normalización tras niveles particularmente elevados en años previos.

En conjunto, el gráfico no solo reafirma que la media global se mantiene estable, sino que también evidencia un ensanchamiento de la brecha entre países y una posible tendencia al estancamiento o deterioro en la felicidad percibida a nivel nacional. La comparación interanual permite detectar trayectorias divergentes y resalta la necesidad de seguir analizando qué factores explican el desempeño diferencial de los países. Esta visualización resulta, por tanto, un complemento clave para comprender la evolución del bienestar en un contexto global marcado por desafíos económicos, políticos y sociales.

Tras analizar la evolución global de la felicidad a nivel de país, resulta interesante centrar la atención en contextos geográficos más acotados que, por su cercanía cultural, económica o política, puedan mostrar dinámicas particulares. En este sentido, se ha seleccionado un grupo de países europeos próximos a España —Portugal, Francia, Italia, Andorra, Suiza, Alemania y Reino Unido— para realizar un análisis comparativo más detallado de la evolución del bienestar subjetivo entre 2015 y 2024.

4.4 Evolución del Happiness Score en España

El objetivo de este bloque es doble: por un lado, examinar la trayectoria específica de España en cuanto a su Happiness Score a lo largo del tiempo; y por otro, contextualizar dicha evolución en relación con sus países vecinos. Además, se analizarán variables complementarias como el PIB per cápita, el apoyo social, la libertad en la toma de decisiones, y las puntuaciones de derechos políticos y libertades civiles, con el fin de identificar posibles factores explicativos de las tendencias observadas.

Este enfoque regional no solo permite evaluar si España sigue dinámicas similares o divergentes respecto a su entorno inmediato, sino que también contribuye a detectar patrones de bienestar y resiliencia en Europa Occidental. Dado el contexto de crisis sucesivas —económica, sanitaria y geopolítica— en el periodo analizado, comparar la evolución de estos países cercanos ofrece una perspectiva enriquecedora sobre el impacto de factores estructurales e institucionales en la percepción de felicidad de la ciudadanía.

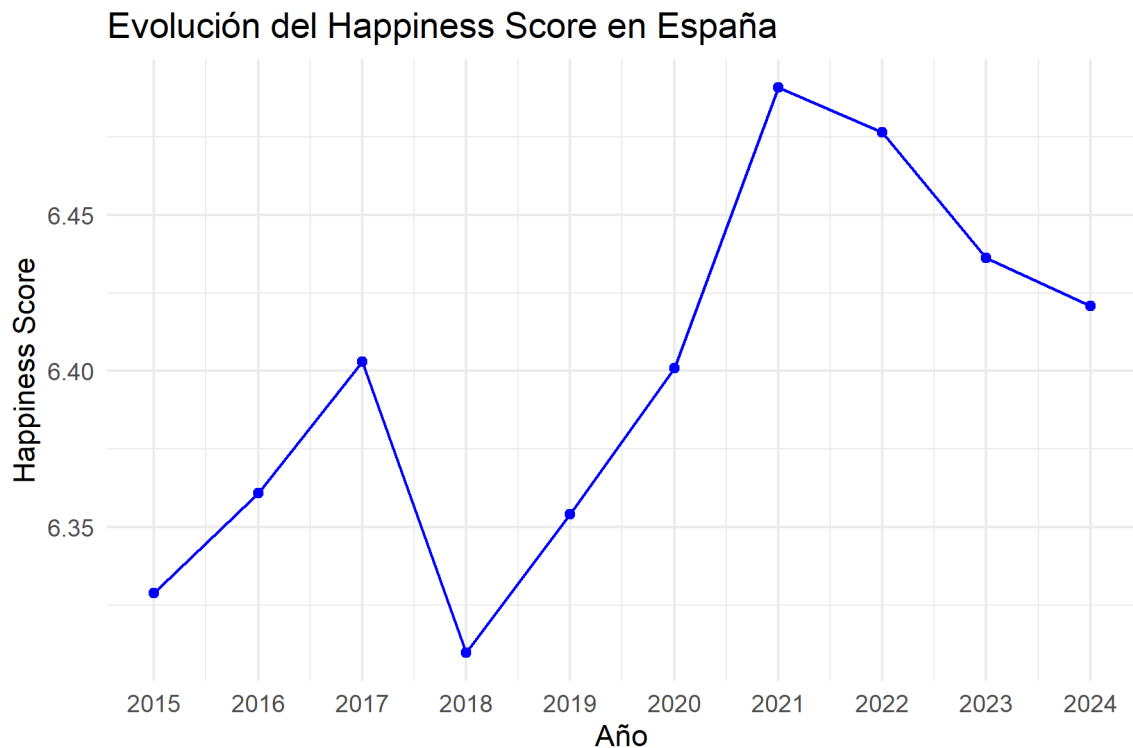


Figura 4.9: Evolución temporal de la felicidad en España

En gráfico de la Figura 4.9, centrado exclusivamente en España, se observa una evolución relativamente estable del índice de felicidad, con pequeñas oscilaciones en torno a valores ligeramente superiores a 6.3. No obstante, destaca una caída en 2018 (mínimo del periodo),

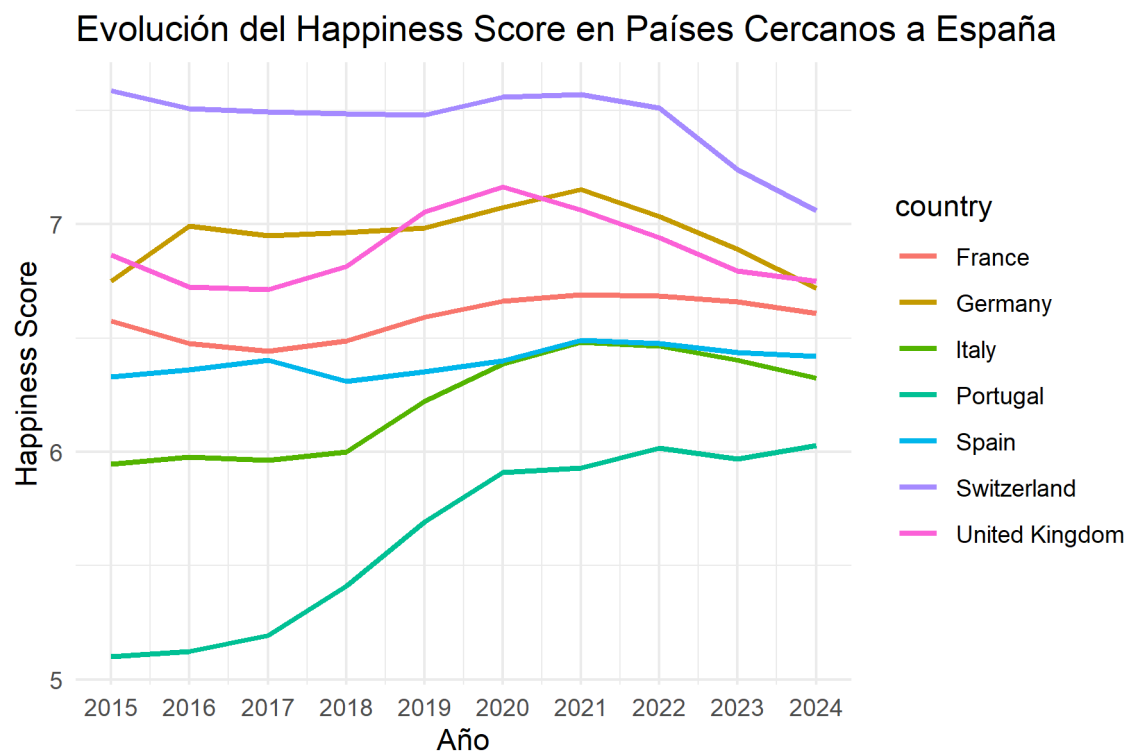


Figura 4.10: Evolución temporal de la felicidad en España y en países cercanos

seguida de una recuperación progresiva que culmina en un pico en 2021, justo tras el primer año completo de pandemia. Este máximo puede parecer contraintuitivo, pero podría explicarse por un efecto de resiliencia colectiva o por cambios en la percepción subjetiva del bienestar ante la crisis sanitaria. Tras 2021, sin embargo, el índice inicia un leve descenso que se prolonga hasta 2024, lo que podría reflejar un efecto retardado de las consecuencias económicas, sociales y emocionales de la pandemia.

El gráfico de la Figura 4.10 contextualiza estos datos al mostrar la evolución del Happiness Score en países próximos a España: Francia, Portugal, Italia, Alemania, Reino Unido y Suiza. Una primera observación relevante es que España se sitúa de forma constante en un rango medio-bajo dentro de este grupo, por encima únicamente de Portugal (hasta 2020) y, ocasionalmente, de Italia. Países como Suiza, Alemania y Reino Unido exhiben sistemáticamente puntuaciones más altas, lo cual podría estar vinculado a mayores niveles de riqueza, estabilidad institucional o servicios sociales más desarrollados.

En términos de trayectoria, destaca el descenso claro en los últimos años (especialmente a partir de 2022) en la mayoría de países analizados, incluyendo aquellos con niveles históricamente altos de felicidad como Suiza o Alemania. Este patrón sugiere la presencia de factores estructurales o globales (crisis energética, inflación, tensiones geopolíticas) que estarían afectando incluso a los países tradicionalmente más satisfechos. En contraste, países como Portugal e Italia muestran una tendencia más estable o incluso creciente hasta 2022, posiblemente como resultado de procesos de recuperación o mejora relativa en relación con sus propias trayectorias históricas.

Los gráficos presentados en la Figura 4.11, Figura 4.12 y Figura 4.13 permiten analizar la evolución de distintas dimensiones del bienestar y la calidad de vida en un conjunto de países europeos cercanos a España entre 2015 y 2024. En conjunto, nos ofrecen una perspectiva comparativa regional que ayuda a contextualizar el caso español en relación con sus vecinos.

En la Figura 4.11, que muestra la evolución del PIB per cápita, se observa una fuerte caída generalizada en 2017, seguida por una recuperación progresiva desde 2018. Esta evolución puede estar relacionada con una recodificación de datos o un cambio metodológico en ese año, pero también con impactos económicos de alcance regional. En cualquier caso, destaca que Suiza mantiene sistemáticamente el mayor PIB per cápita, situándose por encima del resto del grupo. Le siguen Alemania y Reino Unido, con valores elevados y estables. España se encuentra en la parte baja del grupo, junto a Portugal e Italia, aunque todos ellos muestran una tendencia ascendente en los últimos años, lo que refleja una recuperación económica moderada.

En la Figura 4.12, relativo al apoyo social, se aprecia un comportamiento mucho más volátil. En 2018 hay una caída muy pronunciada en todos los países, especialmente en España, Italia y Portugal, lo cual podría reflejar una crisis de confianza institucional o una percepción negativa del entorno social. Sin embargo, en 2019 se produce un repunte, seguido de cierta estabilización con una tendencia ligeramente descendente en algunos países. Suiza y Reino Unido destacan por mantener niveles consistentemente altos de apoyo social, mientras que España mejora en los últimos años tras el fuerte descenso.

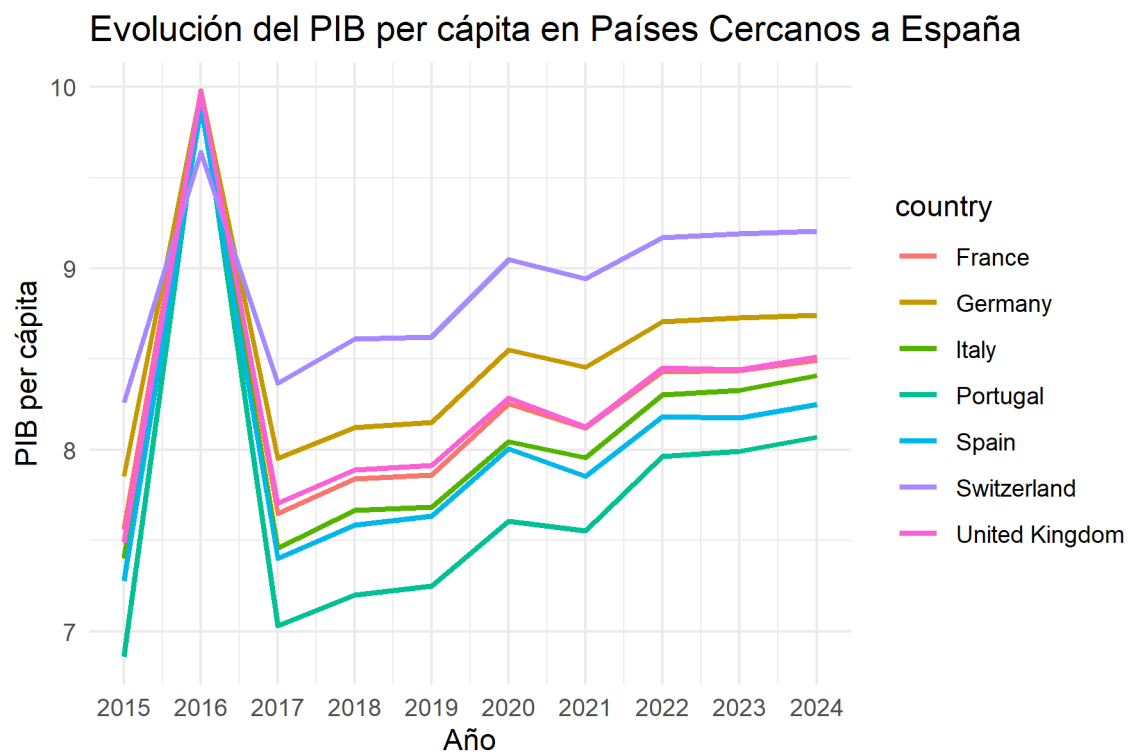


Figura 4.11: Evolución del PIB en España y sus países cercanos

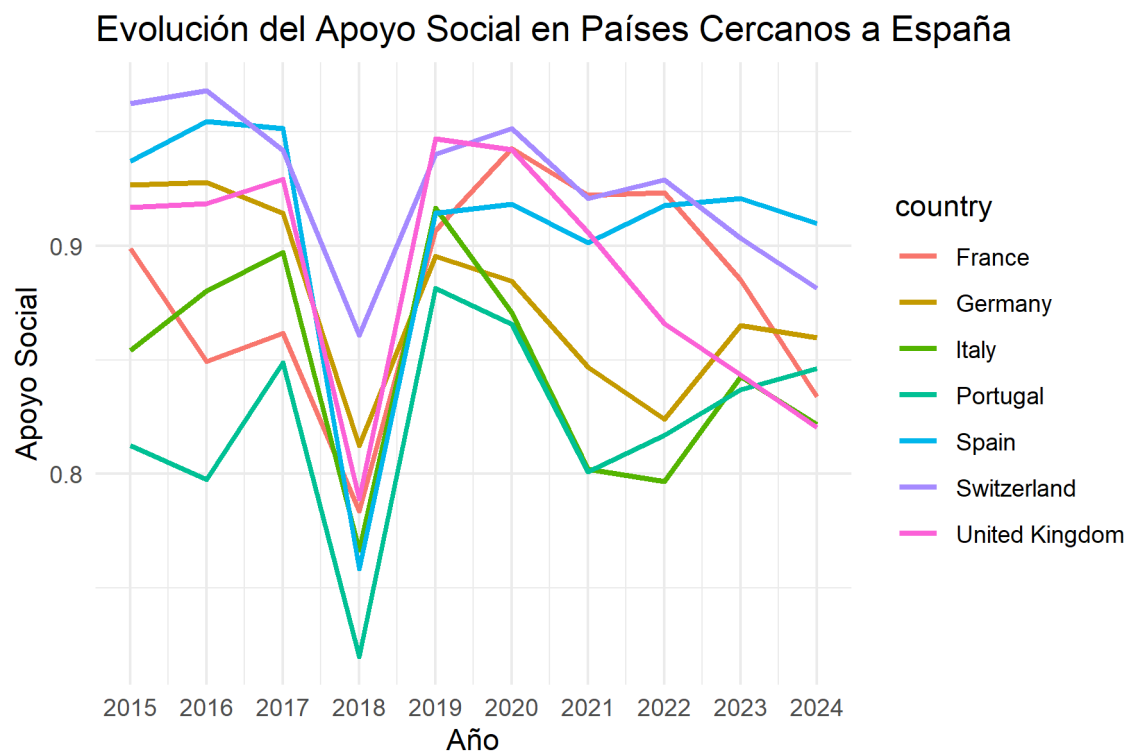


Figura 4.12: Evolución del apoyo social en España y sus países cercanos

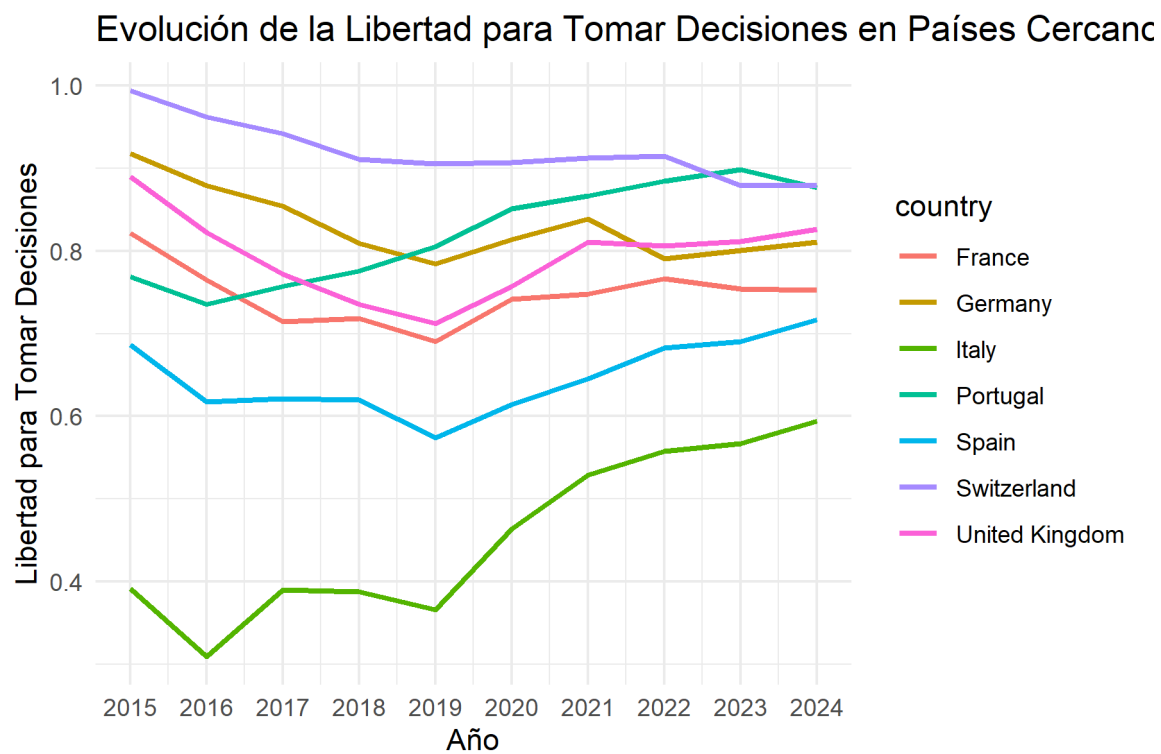


Figura 4.13: Evolución de la libertad en la toma de decisiones en España y sus países cercanos

La Figura 4.13 muestra la evolución de la libertad para tomar decisiones, una dimensión vinculada a la percepción de autonomía personal. Aquí destaca el caso de Suiza, que parte de valores cercanos a 1 y los mantiene prácticamente inalterados. Alemania y Reino Unido también se sitúan en la parte superior, mientras que España parte de un nivel más bajo en 2015 y experimenta una caída hasta 2019, seguida de una recuperación constante a partir de ese momento. Por el contrario, Italia muestra una evolución ascendente sostenida, superando incluso a España en los últimos años. Esta variable presenta un patrón de mejora generalizada en la región, lo que puede reflejar avances en la percepción de libertad individual y gobernanza democrática.

En conjunto, estos gráficos permiten extraer varias conclusiones. En primer lugar, España tiende a situarse en posiciones intermedias o bajas dentro del grupo de comparación, aunque muestra señales de mejora especialmente en PIB y libertad individual. En segundo lugar, países como Suiza y Alemania destacan sistemáticamente por sus buenos resultados, lo que puede atribuirse a sus estructuras institucionales consolidadas y niveles de desarrollo económico. Finalmente, el análisis temporal revela momentos de tensión o crisis (como 2017–2018) seguidos de cierta recuperación, lo que sugiere una resiliencia estructural en la región. Esta comparativa refuerza la utilidad del análisis multinacional para entender mejor los factores que condicionan la evolución del bienestar en contextos políticamente estables pero económicamente diversos.

Aunque al inicio del capítulo ya indicamos que no se detectaron valores atípicos significativos tras el análisis inicial, conviene revisar brevemente algunas observaciones que, por su contexto o comportamiento, podrían considerarse casos especiales. Esta revisión no pretende identificar outliers estadísticos en sentido estricto, sino valorar si existen países cuya evolución difiere notablemente del patrón general y que, por tanto, merecen una atención particular en el análisis interpretativo posterior.

Si observamos la Figura 4.14, de 2015 a 2020 no vemos valores atípicos, pero en 2021 vemos un valor atípico de 2.52, en 2022 de 2.40, en 2023 de 1.86 y en 2024 de 1.72. Todos estos valores atípicos corresponden con Afganistán. Vamos a entrar en más detalle para ver cuál es la evolución de dicho país.

Si observamos la Figura 4.15 y Figura 4.16, Afganistán sufre una gran caída de la felicidad a partir de 2017 y si lo comparamos con países de su región es el que menos puntuación tiene por bastante diferencia. Esto puede haberse dado por diversos factores, como el constante estado de guerra y conflicto en el que se ha encontrado el país, la presencia de los talibanes y otros grupos armados que han aumentado la violencia y el temor en la población, la enorme tasa de pobreza... Vamos a comparar la evolución de Afganistán con sus países vecinos para ver realmente qué puede estar afectando a tal bajada de la felicidad.

Observando la Figura 4.17, vemos que Afganistán presenta el PIB más bajo de la región durante todo el periodo, aunque la diferencia no es tan pronunciada respecto a sus vecinos. La evolución del PIB muestra una tendencia relativamente paralela, lo cual indica que la situación económica general no difiere radicalmente en términos de crecimiento.

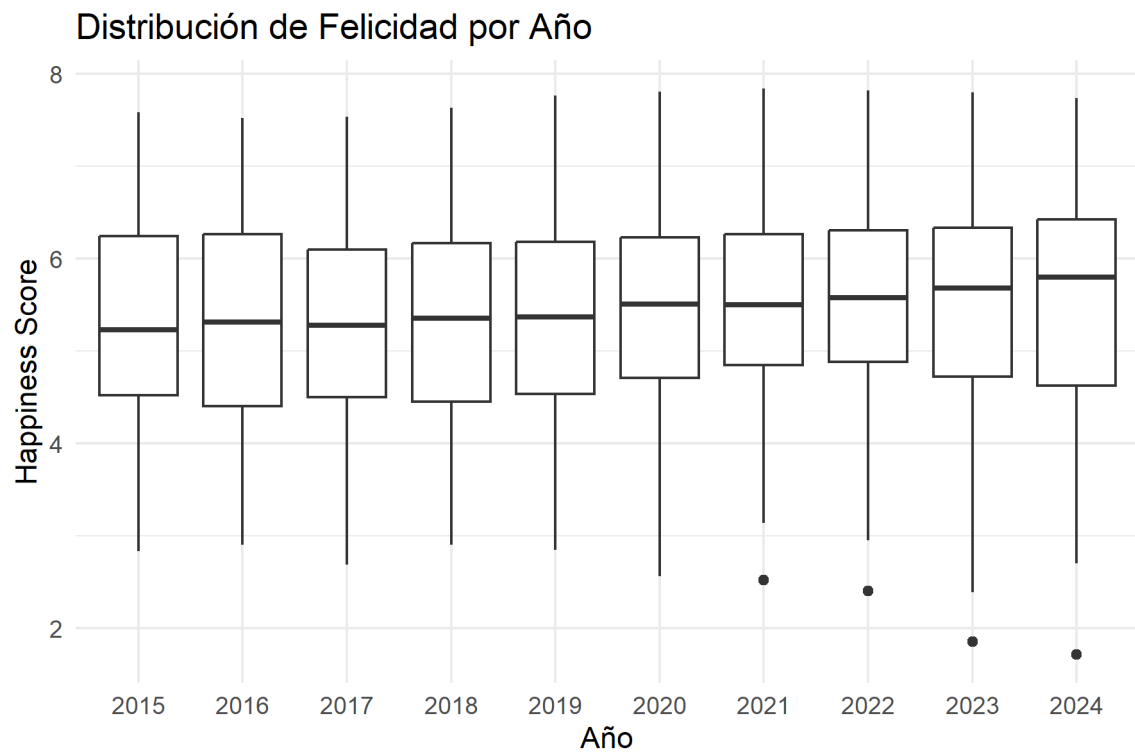


Figura 4.14: Análisis de la distribución de la puntuación de felicidad anual

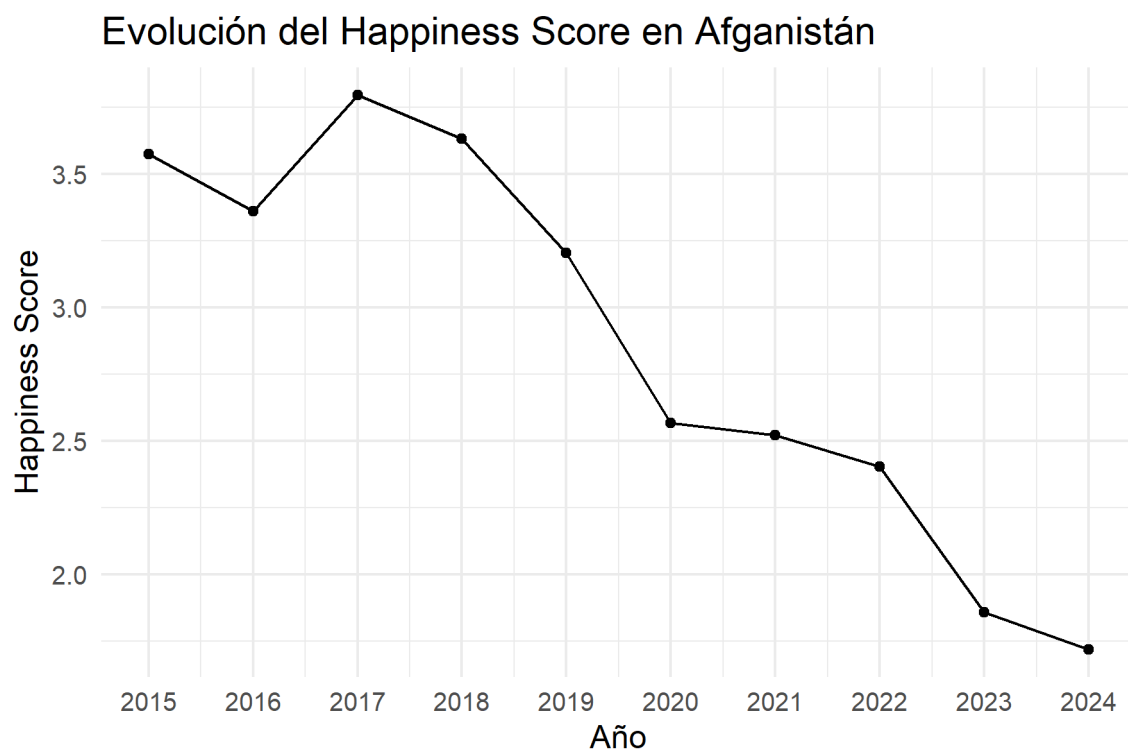


Figura 4.15: Evolución temporal de la felicidad en Afganistán

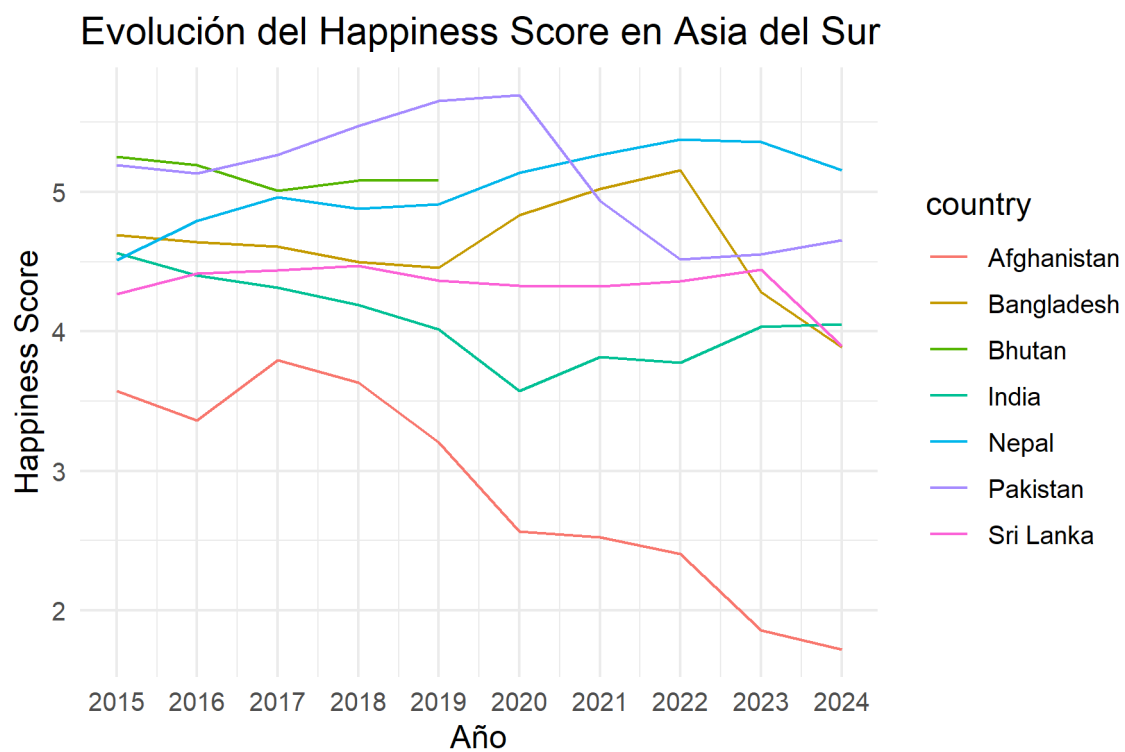


Figura 4.16: Evolución temporal de la felicidad en Afganistán y en países cercanos

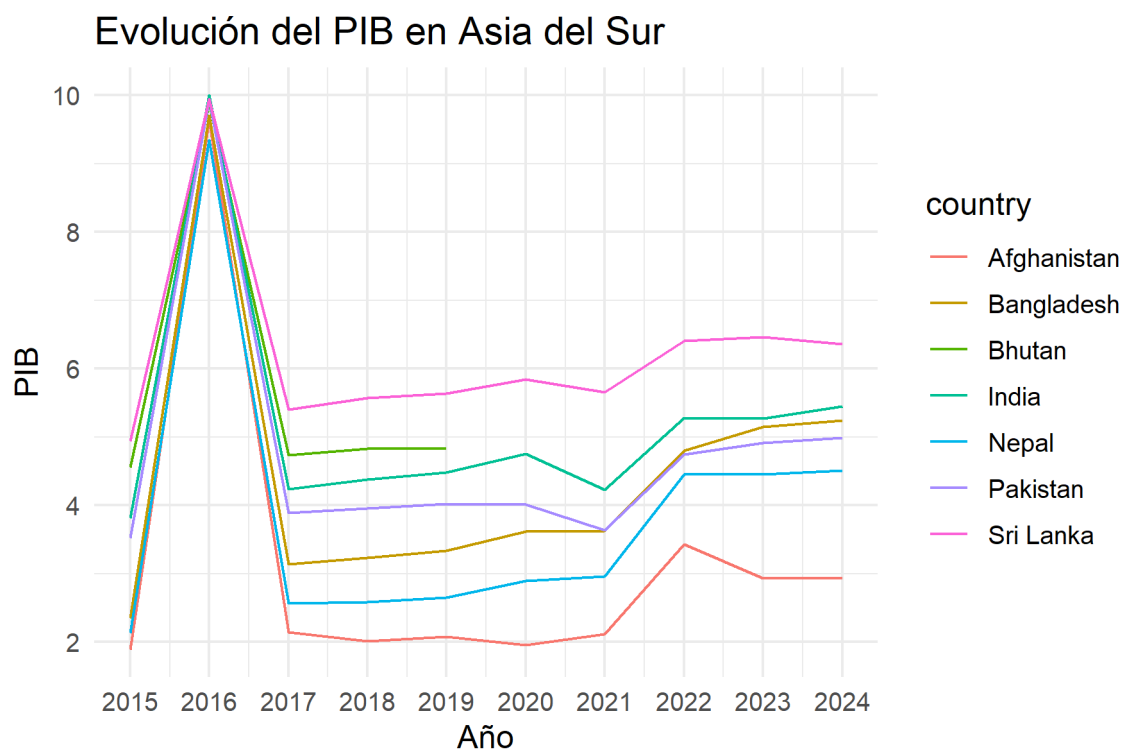


Figura 4.17: Evolución temporal del PIB en Afganistán y en países cercanos

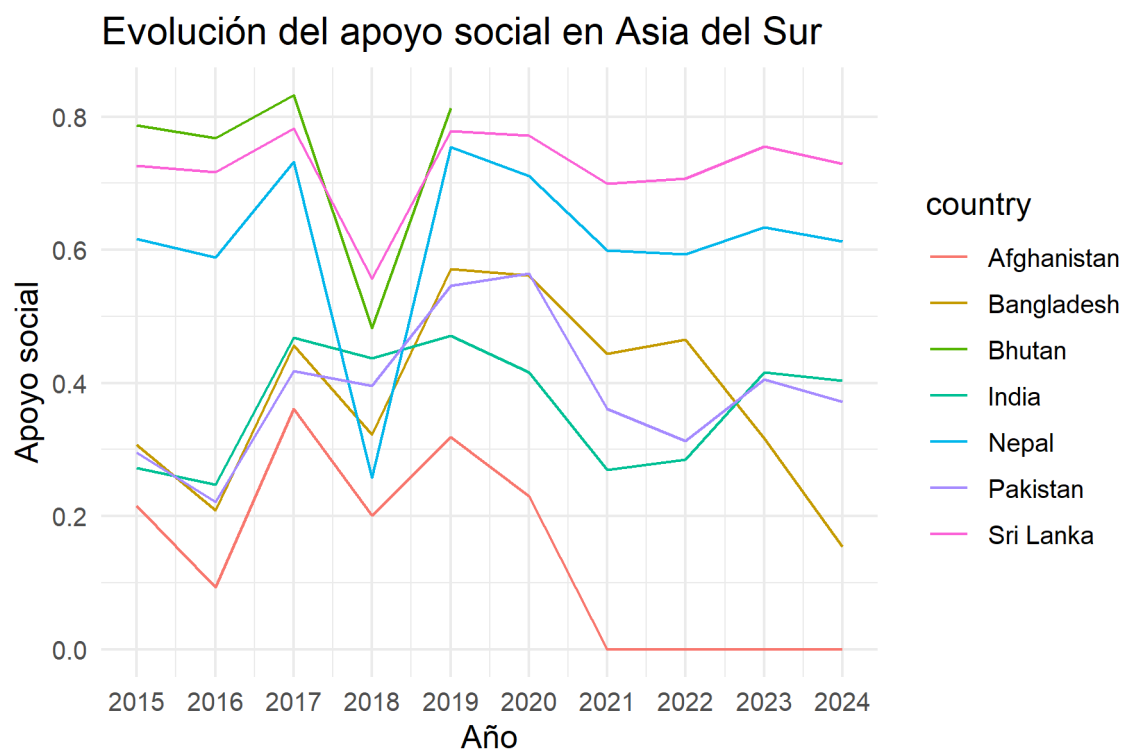


Figura 4.18: Evolución temporal del apoyo social en Afganistán y en países cercanos

Sin embargo, la Figura 4.18 muestra que el apoyo social sufre una caída drástica en Afganistán a partir de 2021, coincidiendo con el regreso de los talibanes al poder. A diferencia de los países vecinos, donde el apoyo se mantiene más estable, en Afganistán prácticamente desaparece, lo que podría explicar un deterioro importante en el bienestar subjetivo de su población.

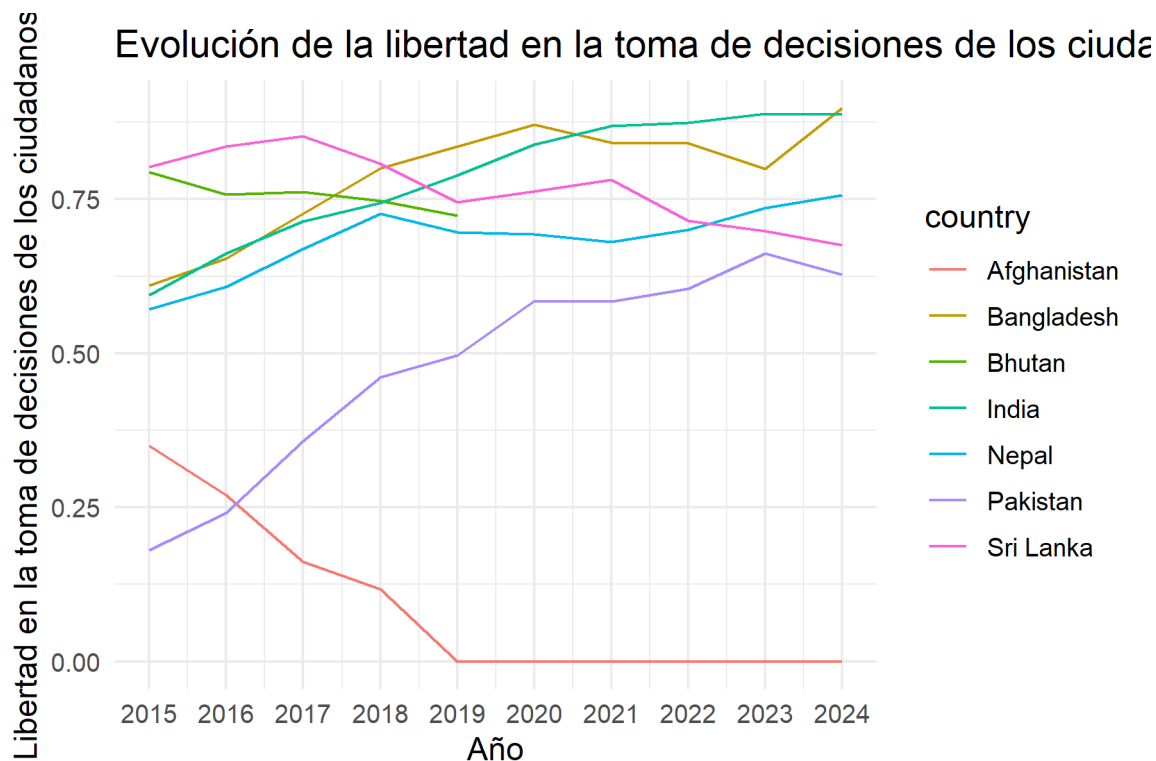


Figura 4.19: Evolución temporal de la libertad de los ciudadanos en Afganistán y en países cercanos

En términos de libertad para tomar decisiones vitales, la Figura 4.19 muestra que Afganistán se sitúa consistentemente por debajo de la media regional. Este valor, que ya era bajo en años anteriores, cae aún más tras 2021, reflejando la pérdida de autonomía percibida por sus ciudadanos.

La Figura 4.20 muestra que la pérdida de libertad se ve reforzada por el deterioro de los derechos políticos, donde Afganistán muestra los valores más altos de privación dentro de la región. El país se convierte en un claro outlier en este aspecto a partir de 2021.

Por último, observando la Figura 4.21, la privación de libertades civiles también coloca a Afganistán en una posición crítica frente a sus vecinos. Este conjunto de restricciones políticas y sociales contribuye directamente a su puntuación tan baja de felicidad, y subraya la necesidad de considerar el contexto institucional al interpretar estos datos.

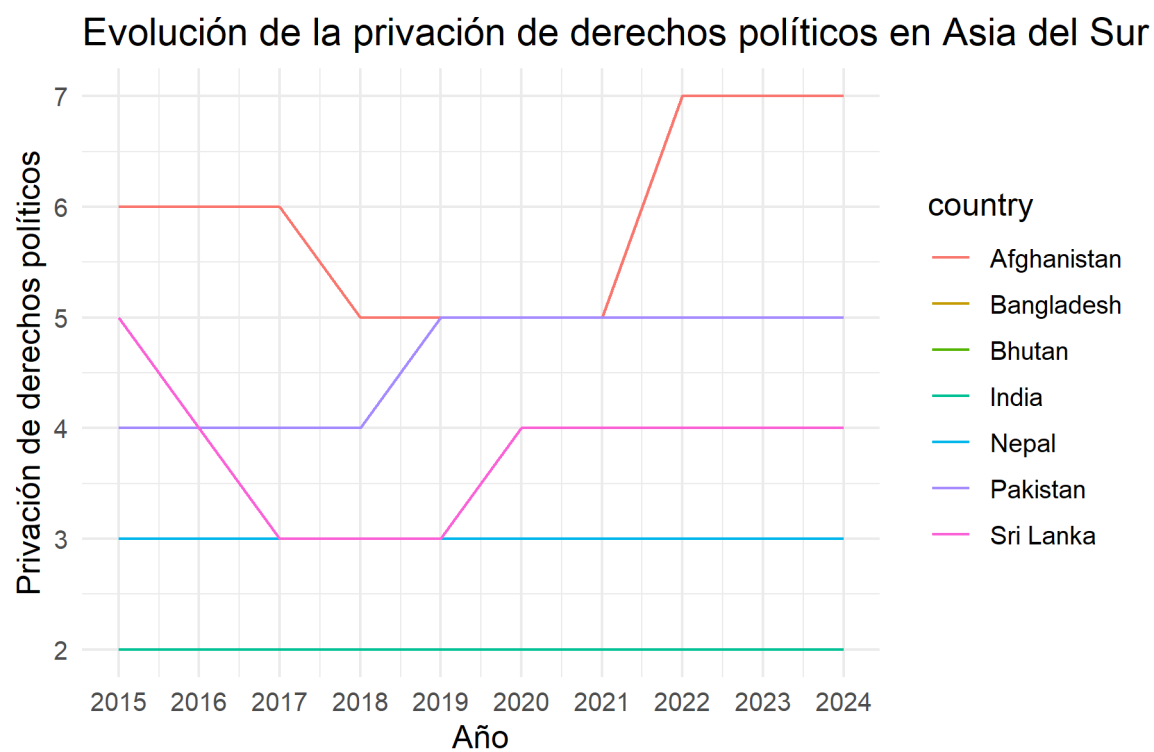


Figura 4.20: Evolución temporal de la privación de derechos políticos en Afganistán y en países cercanos

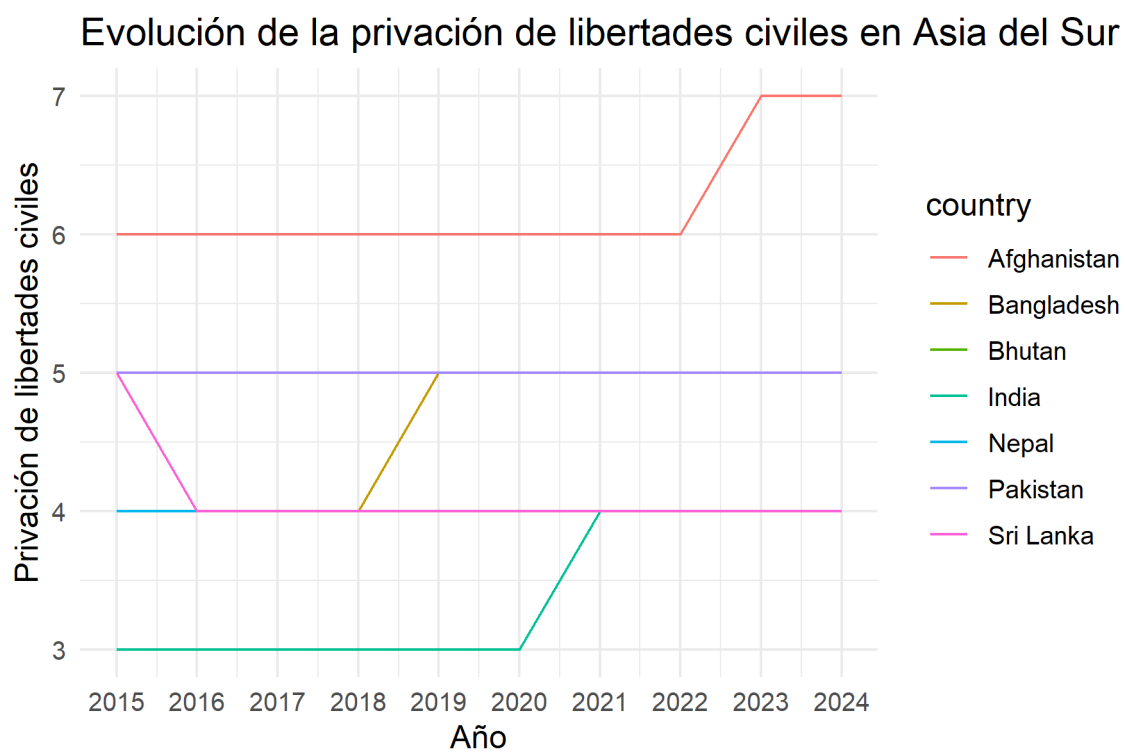


Figura 4.21: Evolución temporal de la privación de libertades civiles en Afganistán y en países cercanos

5 Construcción del modelo

En este capítulo desarrollamos el proceso de construcción de un modelo estadístico para explicar y predecir la percepción de felicidad (Happiness Score) a lo largo del tiempo, utilizando los datos longitudinales disponibles para todos los países entre los años 2015 y 2024. Dado que el análisis de este trabajo se centra en la evolución de la felicidad en el tiempo, consideramos esencial aprovechar toda la información disponible, en lugar de limitar el estudio a una única foto fija (por ejemplo, el año 2024).

El objetivo de este capítulo es construir un modelo que sea estadísticamente sólido y al mismo tiempo interpretable en el contexto de los determinantes sociales, económicos y políticos de la felicidad. Para ello, combinaremos herramientas de la estadística clásica (regresión lineal múltiple) con técnicas específicas para datos longitudinales, incluyendo modelos mixtos que permiten tener en cuenta la estructura jerárquica de los datos (países medidos en diferentes años).

El proceso de modelado se abordará desde dos estrategias complementarias:

- **Top-down:** partimos de un modelo completo que incluye todas las variables relevantes, y vamos eliminando aquellas que no aportan información significativa o que generan problemas como multicolinealidad o sobreajuste.
- **Bottom-up:** comenzamos con un modelo simple con pocas variables y vamos añadiendo progresivamente nuevos predictores, evaluando si su inclusión mejora sustancialmente el ajuste del modelo.

Ambos enfoques nos permiten analizar el trade-off entre complejidad y capacidad explicativa, y encontrar un equilibrio adecuado.

Dado que trabajamos con datos que varían a lo largo del tiempo para cada país, es importante distinguir entre:

- **Variables longitudinales:** cambian con el tiempo (por ejemplo, `gdp`, `support`, `freedom`, `generosity`, `life_exp`, `corruption`).
- **Variables fijas:** son constantes en el tiempo o se utilizan como contexto (por ejemplo, `region`, si se decide mantener como factor).

Además, es importante tener en cuenta la dependencia temporal entre observaciones del mismo país. Este aspecto será considerado al ajustar modelos mixtos con interceptos (y potencialmente pendientes) aleatorios por país.

5.1 Análisis exploratorio y selección inicial de variables

Aunque ya se analizó previamente la estructura de los datos, antes de ajustar cualquier modelo es útil revisar de nuevo las características más relevantes desde una perspectiva predictiva, identificando posibles problemas (como valores faltantes o outliers) y seleccionando las variables que podrían actuar como buenos predictores del `happiness_score`.

5.1.1 Estructura del dataset longitudinal

Nuestro conjunto de datos contiene observaciones anuales de múltiples países en el período 2015–2024. Para cada país y año, disponemos de una serie de variables socioeconómicas y de percepción ciudadana. Las variables disponibles son:

- `happiness_score`: puntuación de felicidad (variable respuesta)
- `gdp`: producto interior bruto per cápita
- `support`: percepción de apoyo social
- `life_exp`: esperanza de vida saludable
- `freedom`: libertad para tomar decisiones
- `generosity`: generosidad de la población
- `corruption`: percepción de corrupción en instituciones
- `country`: identificador del país
- `year`: año de la observación
- `regional_indicator`: región a la que pertenece el país

Dado el enfoque longitudinal, usaremos `country` como unidad de agrupación para modelar efectos específicos de cada país a lo largo del tiempo.

5.1.2 Visualización y evolución temporal de las variables

A continuación, exploramos cómo varían las principales variables predictoras a lo largo del tiempo. Esto nos ayuda a evaluar la estabilidad o variabilidad de cada predictor y su potencial relación con el `happiness_score`.

Observando la Figura 5.1, vemos que la esperanza de vida muestra una tendencia general al alza, con ciertas oscilaciones intermedias que podrían estar asociadas a eventos sanitarios globales como la pandemia de COVID-19. El incremento observado en 2022 es particularmente destacado y podría requerir un análisis más desagregado por región. Las variables sociales de la Figura 5.2 presentan comportamientos diferenciados. La percepción de corrupción (`corruption`) muestra variaciones abruptas que podrían reflejar problemas de medición o eventos políticos clave. La libertad de tomar decisiones (`freedom`) y el apoyo social (`support`) se mantienen relativamente estables, mientras que la generosidad (`generosity`) parece haber aumentado gradualmente en los últimos años. La evolución del PIB (`gdp`) en la

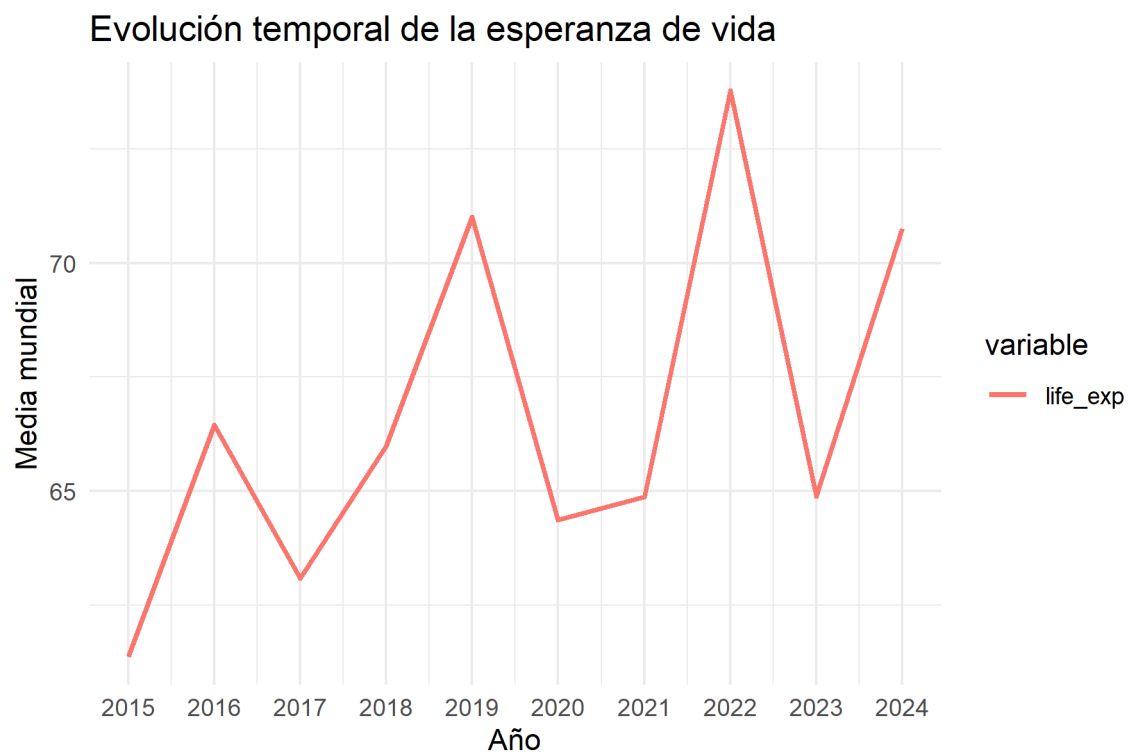


Figura 5.1: Evolución temporal de la esperanza de vida

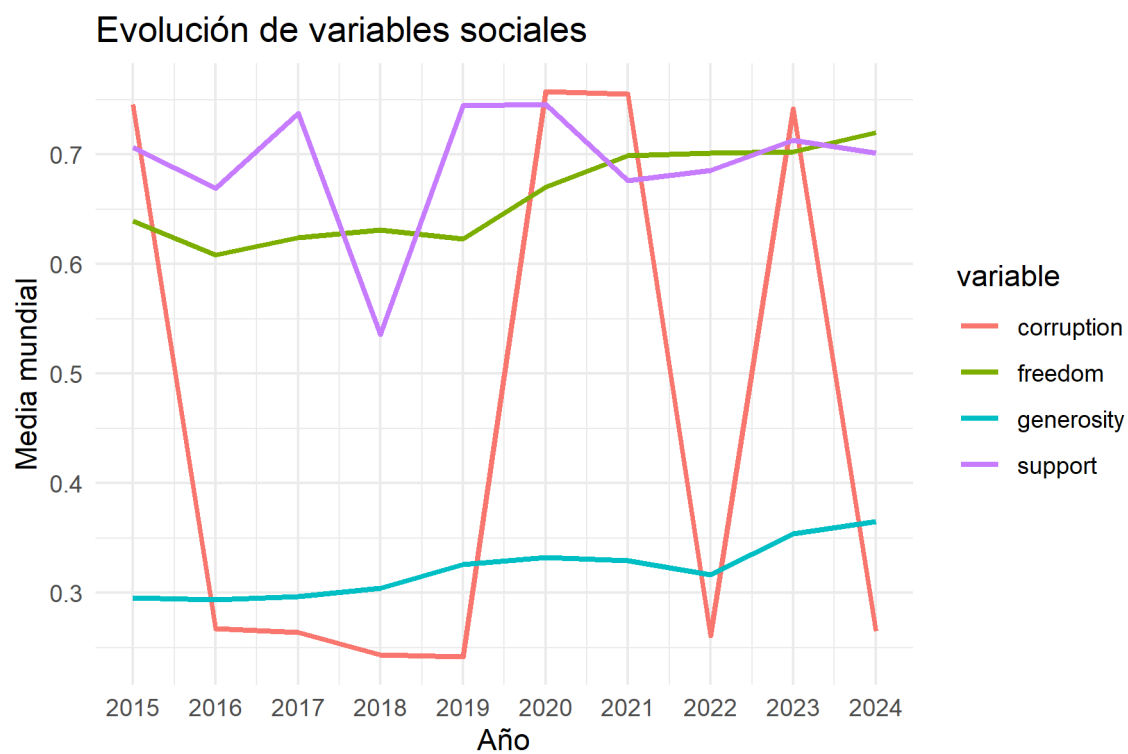


Figura 5.2: Evolución temporal de la percepción de corrupción, libertad, generosidad y apoyo social

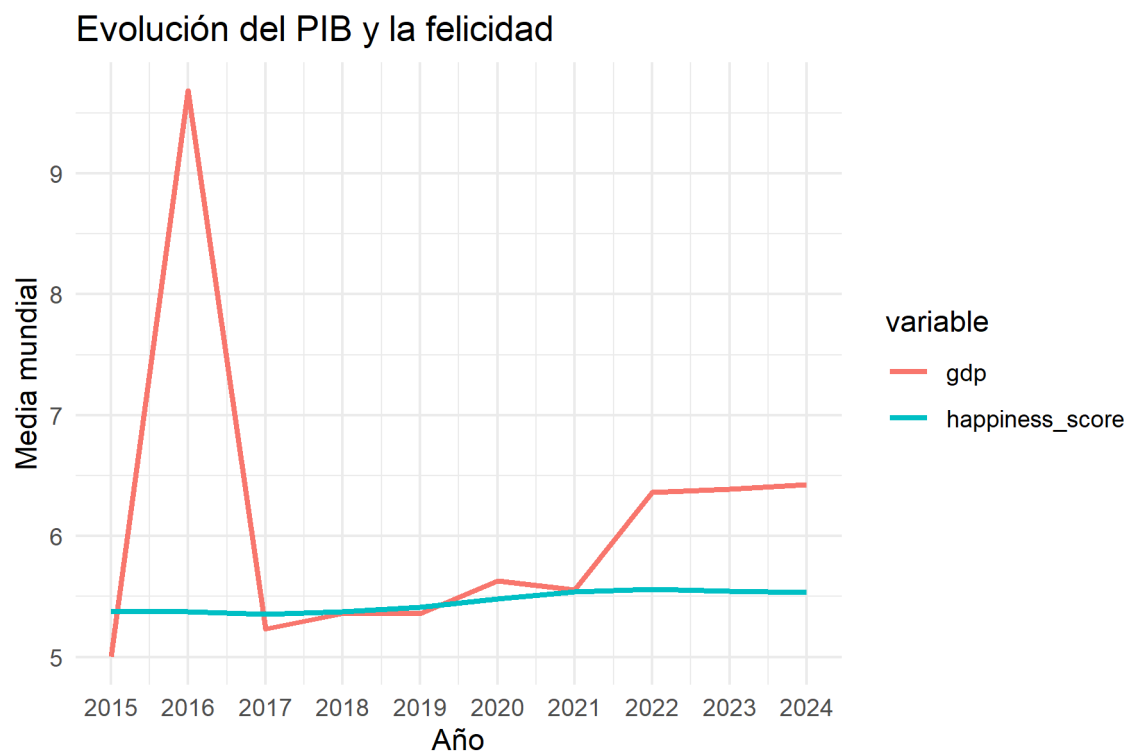


Figura 5.3: Evolución temporal del PIB y la felicidad

Figura 5.3 muestra un aumento importante entre 2020 y 2022, que puede deberse a efectos de recuperación tras la pandemia. El `happiness_score` permanece mucho más estable a lo largo del tiempo, lo que sugiere que el crecimiento económico no se traduce automáticamente en un incremento proporcional de la felicidad percibida. Sin embargo, también podría interpretarse como un efecto con retardo, en el que los beneficios del aumento del PIB se reflejan en el bienestar subjetivo con cierto desfase temporal.

5.1.3 Matriz de correlaciones

Antes de ajustar ningún modelo, es útil examinar la correlación entre las variables explicativas y la variable objetivo `happiness_score`. Esto nos permitirá identificar posibles relaciones lineales, evaluar redundancias y tomar decisiones informadas sobre qué variables incluir inicialmente en el modelo.

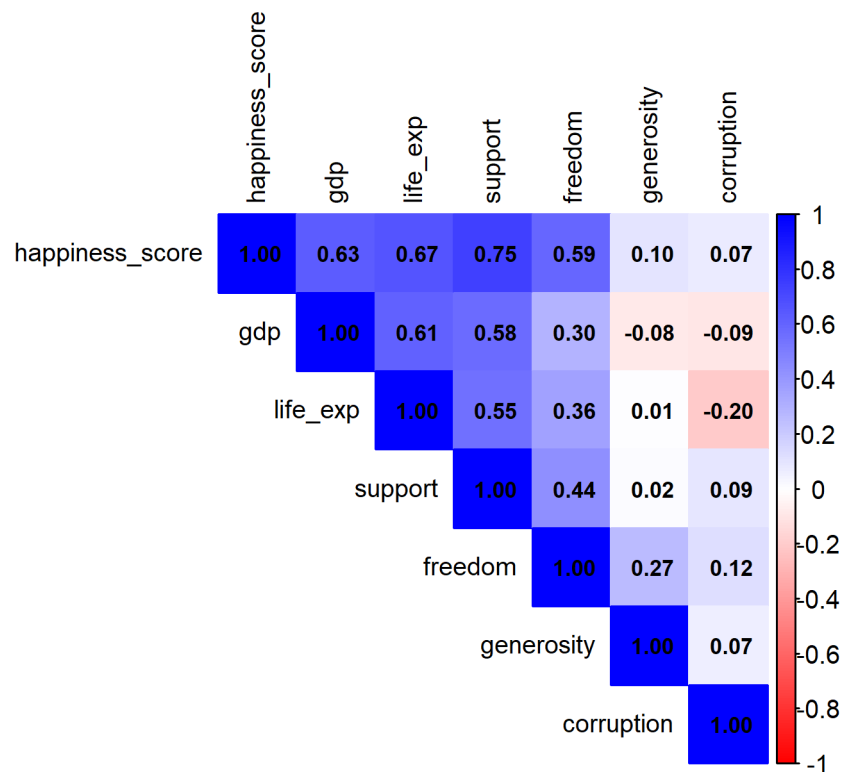


Figura 5.4: Matriz de correlaciones de las variables de nuestra base de datos

Como se puede observar en la Figura 5.4, las correlaciones más fuertes con `happiness_score` corresponden a:

- support (0.75): el apoyo social es la variable que más se relaciona con la felicidad percibida.

- life_exp (0.67) y gdp (0.63): muestran relaciones importantes, lo que refleja la relevancia del bienestar económico y la salud.
- freedom (0.59) también presenta una relación moderada.

En cambio, variables como generosity (0.10) y corruption (0.07) muestran una relación muy débil con la felicidad, lo que hace cuestionable su relevancia explicativa. Las correlaciones entre predictores no son excesivamente altas (ninguna supera 0.8), por lo que, en principio, no se anticipan grandes problemas de multicolinealidad. Aun así, esto se verificará más adelante con el cálculo de los VIF (Variance Inflation Factor).

Este análisis justifica incluir support, life_exp, gdp y freedom en una primera versión del modelo, aunque se explorarán también modelos más parciales para comparar rendimiento y parsimonia.

5.2 Criterios de selección del modelo

Al construir un modelo estadístico, es común que existan múltiples combinaciones posibles de predictores. Para elegir la especificación más adecuada, se utilizan criterios que balancean dos aspectos fundamentales: el ajuste al conjunto de datos (qué tan bien predice el modelo los datos observados), y la complejidad del modelo (cuántos parámetros se incluyen). Dos de los criterios más utilizados para este propósito son el Akaike Information Criterion (AIC) y el Bayesian Information Criterion (BIC).

AIC penaliza la complejidad del modelo y busca minimizar la pérdida de información. Se calcula como: Dos de los criterios más utilizados para este propósito son el Akaike Information Criterion (AIC) y el Bayesian Information Criterion (BIC).

AIC penaliza la complejidad del modelo y busca minimizar la pérdida de información. Se calcula como:

$$AIC = -2 \cdot \log(\hat{L}) + 2k$$

donde \hat{L} es la verosimilitud máxima del modelo y k el número de parámetros.

BIC penaliza más fuertemente los modelos complejos (dependiendo del tamaño de muestra n):

$$BIC = -2 \cdot \log(\hat{L}) + k \cdot \log(n)$$

Un AIC o BIC más bajo indica un mejor modelo, pero el AIC tiende a seleccionar modelos más complejos (es más permisivo); mientras que el BIC favorece modelos más parsimoniosos (penaliza más la complejidad). Ambos criterios se pueden usar para comparar modelos con los mismos datos y la misma variable respuesta.

5.3 Modelado clásico

Como punto de partida, construimos un modelo clásico de regresión lineal múltiple para explicar el `happiness_score` a partir de variables como `gdp`, `life_exp`, `support`, `freedom`, `generosity` y `corruption`. Para explorar la mejor combinación de predictores, aplicamos dos estrategias de selección de variables:

5.3.1 Estrategia top-down (backward elimination)

Usamos el criterio AIC y BIC para eliminar aquellas variables cuya eliminación mejora la simplicidad del modelo sin sacrificar capacidad predictiva.

Start: AIC=-1590.33

```
happiness_score ~ gdp + life_exp + support + freedom + generosity +  
corruption
```

	Df	Sum of Sq	RSS	AIC
<none>			496.37	-1590.3
- generosity	1	2.207	498.57	-1585.8
- corruption	1	10.559	506.93	-1561.3
- gdp	1	35.828	532.19	-1489.6
- life_exp	1	73.315	569.68	-1389.3
- freedom	1	88.735	585.10	-1349.9
- support	1	132.429	628.80	-1243.7

Start: AIC=-1553.26

```
happiness_score ~ gdp + life_exp + support + freedom + generosity +  
corruption
```

	Df	Sum of Sq	RSS	AIC
- generosity	1	2.207	498.57	-1554.0
<none>			496.37	-1553.3
- corruption	1	10.559	506.93	-1529.5
- gdp	1	35.828	532.19	-1457.8
- life_exp	1	73.315	569.68	-1357.5
- freedom	1	88.735	585.10	-1318.1
- support	1	132.429	628.80	-1212.0

Step: AIC=-1554.02

```
happiness_score ~ gdp + life_exp + support + freedom + corruption
```

	Df	Sum of Sq	RSS	AIC
<none>			498.57	-1554.0
- corruption	1	10.754	509.33	-1529.9
- gdp	1	34.143	532.72	-1463.7
- life_exp	1	73.635	572.21	-1358.3
- freedom	1	106.041	604.61	-1277.1
- support	1	131.548	630.12	-1216.2

En la selección según AIC, el modelo presenta un AIC final de -1590.3, donde la variable generosity fue descartada por no aportar mejora significativa al modelo, lo cual concuerda con su baja correlación con la variable objetivo observada anteriormente. Por su parte, el BIC penaliza más severamente los modelos complejos, favoreciendo modelos más parsimoniosos. En este caso, el modelo óptimo según BIC también descarta generosity, y produce exactamente la misma especificación final que el modelo por AIC. Esto refuerza la robustez del modelo seleccionado, al coincidir ambos criterios a pesar de su diferencia en la penalización por número de parámetros.

5.3.2 Estrategia bottom-up (forward selection)

También exploramos una estrategia bottom-up (selección hacia adelante), que parte de un modelo nulo y añade variables una a una en función de la mejora del AIC. Esta estrategia permite comprobar si existe alguna combinación alternativa de predictores que produzca un modelo competitivo o incluso mejor al obtenido por eliminación hacia atrás.

Start: AIC=369.95
happiness_score ~ 1

	Df	Sum of Sq	RSS	AIC
+ support	1	1052.77	839.17	-826.33
+ life_exp	1	846.37	1045.57	-502.19
+ gdp	1	762.22	1129.72	-388.10
+ freedom	1	661.35	1230.60	-262.03
+ generosity	1	20.45	1871.50	355.93
+ corruption	1	10.22	1881.72	363.96
<none>			1891.95	369.95

Step: AIC=-826.33
happiness_score ~ support

	Df	Sum of Sq	RSS	AIC
+ life_exp	1	179.860	659.31	-1179.89

+ freedom	1	165.587	673.59	-1148.32
+ gdp	1	116.925	722.25	-1045.50
+ generosity	1	15.313	823.86	-851.48
<none>			839.17	-826.33
+ corruption	1	0.023	839.15	-824.37

Step: AIC=-1179.89

happiness_score ~ support + life_exp

	Df	Sum of Sq	RSS	AIC
+ freedom	1	117.915	541.40	-1468.3
+ gdp	1	31.208	628.11	-1249.4
+ corruption	1	20.373	638.94	-1224.2
+ generosity	1	15.743	643.57	-1213.5
<none>			659.31	-1179.9

Step: AIC=-1468.33

happiness_score ~ support + life_exp + freedom

	Df	Sum of Sq	RSS	AIC
+ gdp	1	32.072	509.33	-1556.3
+ corruption	1	8.683	532.72	-1490.2
<none>			541.40	-1468.3
+ generosity	1	0.642	540.76	-1468.1

Step: AIC=-1556.34

happiness_score ~ support + life_exp + freedom + gdp

	Df	Sum of Sq	RSS	AIC
+ corruption	1	10.7540	498.57	-1585.8
+ generosity	1	2.4018	506.93	-1561.3
<none>			509.33	-1556.3

Step: AIC=-1585.79

happiness_score ~ support + life_exp + freedom + gdp + corruption

	Df	Sum of Sq	RSS	AIC
+ generosity	1	2.2067	496.37	-1590.3
<none>			498.57	-1585.8

Step: AIC=-1590.33

happiness_score ~ support + life_exp + freedom + gdp + corruption +
generosity

En este caso, el procedimiento fue incorporando progresivamente variables hasta alcanzar el modelo completo, es decir, con todos los predictores. Este resultado contrasta con el modelo obtenido por backward elimination, en el que algunas variables eran descartadas. Esta diferencia refleja que los criterios de selección pueden llevar a soluciones distintas según el punto de partida.

En base a los resultados anteriores, el modelo final que utilizaremos es: $\text{happiness_score} \sim \text{gdp} + \text{life_exp} + \text{support} + \text{freedom} + \text{corruption}$. Este modelo muestra un equilibrio adecuado entre capacidad predictiva y parsimonia. Como veremos en la siguiente sección, se procederá ahora a evaluar el ajuste del modelo, su interpretación y su validación diagnóstica.

5.3.3 Diagnóstico y validación final del modelo

Una vez seleccionado el modelo de regresión múltiple más adecuado según los criterios AIC y BIC, es fundamental verificar si cumple con las hipótesis necesarias para garantizar su validez estadística. Estas hipótesis incluyen:

- Normalidad de los residuos.
- Media cero de los residuos.
- Homoscedasticidad (varianza constante de los errores).
- Independencia de los errores.

5.3.3.1 Normalidad de los residuos

Se aplica el test de Jarque-Bera para evaluar si los residuos del modelo siguen una distribución normal. El resultado indica un p-valor muy bajo, lo que nos lleva a rechazar la hipótesis de normalidad.

El p-valor obtenido es extremadamente bajo ($p\text{-value} = 2.195e-10$), lo que indica que los residuos no siguen una distribución normal. Esto también se confirma visualmente en el gráfico Q-Q de la Figura 5.5, donde los residuos llegan un momento en el que se desvían de la línea teórica.

5.3.3.2 Media cero

Calculamos la media de los residuos para verificar si se aproxima a cero, como requiere el modelo. En este caso, se cumple adecuadamente:

Si observamos la gráfica Residuals vs Fitted en la Figura 5.5, podríamos decir que los residuos están uniformemente dispersos alrededor del eje de abscisas, por lo que se cumple la condición de que los residuos tienen media cero.

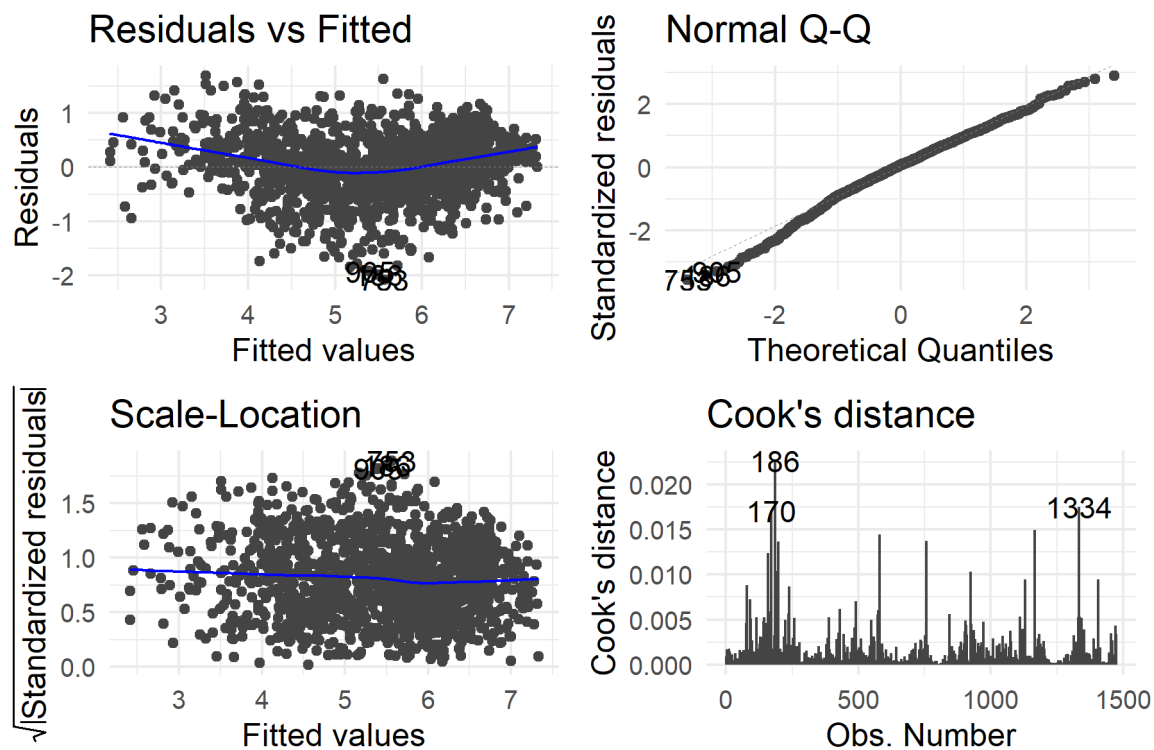


Figura 5.5: Gráficas para la validación del modelo clásico

5.3.3.3 Homoscedasticidad

Se aplica la prueba de Breusch-Pagan para comprobar si la varianza de los errores es constante. El resultado del test devuelve un p-valor bajo, lo que sugiere la existencia de heterocedasticidad.

El gráfico de residuos frente a los valores ajustados de la Figura 5.5 también muestra un patrón creciente, indicio visual de heterocedasticidad. El test devuelve un p-valor $< 2.2e-16$, por lo que rechazamos la hipótesis nula de homocedasticidad. Esto indica que hay heterocedasticidad en los residuos, es decir, su varianza no es constante.

5.3.3.4 No correlación de los errores

Mediante el test de Durbin-Watson, se verifica que no exista autocorrelación en los residuos. En este caso, el p-valor es suficientemente alto como para no rechazar la hipótesis nula, por lo que la independencia de los errores se mantiene.

Con un estadístico de Durbin-Watson de 1.4461 y un p-valor $< 2.2e-16$, se concluye que existe autocorrelación positiva entre los residuos.

5.3.3.5 Conclusión del diagnóstico

Estos incumplimientos sugieren que el modelo, aunque aparentemente ajustado, no es estadísticamente válido en un contexto longitudinal. Específicamente, ignora la dependencia entre observaciones del mismo país a lo largo del tiempo, lo que puede sesgar los resultados.

La inadecuación del modelo clásico justifica el uso de enfoques más robustos, capaces de incorporar la estructura jerárquica de los datos. En particular, los modelos lineales mixtos (LMM) permiten modelar efectos aleatorios por país, capturar la correlación intrínseca entre medidas repetidas y mejorar la validez estadística y la interpretación de los resultados. En la siguiente sección abordamos su aplicación utilizando la función `lmer()` del paquete `lme4`.

5.4 Modelos Lineales Mixtos (LMM)

En el análisis clásico mediante regresión lineal múltiple, se asumía que las observaciones eran independientes entre sí. Sin embargo, en nuestro caso trabajamos con datos longitudinales, es decir, con observaciones repetidas a lo largo del tiempo para los mismos países. Esto introduce una dependencia entre observaciones dentro del mismo país, que los modelos clásicos no pueden capturar adecuadamente.

Para abordar esta limitación, recurrimos a los modelos lineales mixtos (LMM). Estos modelos permiten combinar efectos fijos, que capturan el efecto promedio de los predictores sobre

la variable respuesta en toda la población; y efectos aleatorios, que permiten modelar la variabilidad específica entre países, incorporando interceptos (y potencialmente pendientes) distintos para cada uno. De esta forma podemos capturar la estructura jerárquica de los datos, reconociendo que cada país puede tener un nivel base de felicidad distinto (intercepto propio), mientras que los efectos de las variables predictoras son comunes a todos los países. En nuestro modelo, se consideran efectos fijos aquellas variables explicativas cuyo efecto queremos estimar de forma generalizable para toda la población (en este caso, todos los países y años). Estas variables incluyen gdp, support, freedom, life_exp, corruption y las variables políticas (is_democracy, regime_category, etc.). Se introduce un efecto aleatorio por país (intercepto aleatorio), porque cada país tiene un nivel base distinto de felicidad no explicado por las variables fijas, hay dependencia entre observaciones del mismo país en distintos años, y no nos interesa estimar el efecto específico de cada país, sino tener en cuenta la variabilidad entre ellos. En nuestro caso, el modelo lineal mixto adoptado es el siguiente:

$$\begin{aligned} \text{happiness}_{ij} = & \beta_0 + \beta_1 \cdot \text{support}_{ij} + \beta_2 \cdot \text{lifeexp}_{ij} + \beta_3 \cdot \text{freedom}_{ij} \\ & + \beta_4 \cdot \text{gdp}_{ij} + \beta_5 \cdot \text{corruption}_{ij} + \beta_6 \cdot \text{generosity}_{ij} + \beta_7 \cdot \text{year}_{ij} \\ & + u_{0j} + \varepsilon_{ij} \end{aligned}$$

donde:

- happiness_{ij} es la puntuación de felicidad observada para el país j en el año i .
- **support**, **life_exp**, **freedom**, **gdp**, **corruption**, **generosity** y **year** se incluyen como efectos fijos, comunes a todos los países.
- β_0, \dots, β_6 son los coeficientes de estos efectos fijos.
- $u_{0j} \sim \mathcal{N}(0, \sigma_u^2)$ representa un efecto aleatorio de intercepto para el país j , permitiendo variación en la felicidad media entre países.
- $\varepsilon_{ij} \sim \mathcal{N}(0, \sigma^2)$ es el término de error individual, independiente e idénticamente distribuido.

```
Linear mixed model fit by REML ['lmerMod']
```

```
Formula: happiness_score ~ support + life_exp + freedom + gdp + corruption +  
generosity + year + (1 | country)
```

```
Data: df_sin_democracia_sin_na
```

```
REML criterion at convergence: 1314.9
```

```
Scaled residuals:
```

	Min	1Q	Median	3Q	Max
	-4.8206	-0.4861	0.0061	0.5430	3.9844

Random effects:

Groups	Name	Variance	Std.Dev.
country	(Intercept)	0.5824	0.7631
Residual		0.0895	0.2992

Number of obs: 1474, groups: country, 160

Fixed effects:

	Estimate	Std. Error	t value
(Intercept)	15.144550	6.970186	2.173
support	0.783273	0.081751	9.581
life_exp	0.010629	0.002473	4.298
freedom	1.179542	0.098470	11.979
gdp	0.012509	0.005411	2.312
corruption	0.048587	0.029645	1.639
generosity	0.351233	0.131270	2.676
year	-0.005920	0.003501	-1.691

Correlation of Fixed Effects:

	(Intr)	supprt	lif_xp	freedm	gdp	crrptn	gnrsty
support	-0.043						
life_exp	0.466	-0.015					
freedom	0.238	-0.156	0.001				
gdp	-0.071	-0.014	-0.066	-0.041			
corruption	0.223	-0.107	0.511	-0.090	0.151		
generosity	0.192	0.006	-0.051	-0.221	0.089	-0.056	
year	-1.000	0.037	-0.483	-0.241	0.067	-0.232	-0.192

Si nos fijamos en los efectos aleatorios, la desviación estándar del intercepto por país es de 0.7631, lo que refleja una importante variabilidad estructural entre países; mientras que el término residual tiene desviación estándar 0.2992, menor que la del intercepto aleatorio, lo que refuerza la utilidad del modelo mixto. Analizando los efectos fijos, vemos como todas las variables menos corruption y year son significativas; pero las variables más destacadas son support, generosity y freedom ya que son aquellas que, debido al alto valor de sus coeficientes estimados, mayor impacto tienen en el modelo.

Antes de proceder al diagnóstico exhaustivo del modelo propuesto, es importante evaluar si esta especificación es la más adecuada para los datos. Para ello, se comparan múltiples modelos candidatos que varían en los predictores incluidos, utilizando el criterio de información de Akaike (AIC), previamente descrito. En el caso de modelos mixtos, existe una distinción importante entre dos métodos de estimación:

- REML (Restricted Maximum Likelihood): se recomienda cuando el objetivo es obtener estimaciones precisas de la varianza de los efectos aleatorios, manteniendo fija la

estructura de efectos fijos. No se debe usar REML para comparar modelos con diferentes predictores fijos.

- ML (Maximum Likelihood): se utiliza para comparar modelos que difieren en los efectos fijos (como sucede en selección de variables), ya que estima la verosimilitud sin condicionar por los efectos fijos.

Por este motivo, para aplicar procedimientos de selección, reajustamos el modelo mixto utilizando ML, y comparamos los AIC de cada especificación candidata. Para identificar el modelo que mejor equilibra capacidad explicativa y simplicidad, utilizamos la función `dredge()` del paquete MuMIn. Esta herramienta realiza una búsqueda exhaustiva de modelos posibles a partir de un modelo base, generando todas las combinaciones de predictores (con y sin interacción si se desea), y calcula para cada uno de ellos métricas como el AIC, AICc o BIC. El objetivo es seleccionar el modelo que tenga el menor valor de AICc, lo que indica el mejor compromiso entre ajuste y complejidad.

El funcionamiento de `dredge()` parte del modelo saturado (especificado con todos los predictores que se desean considerar) y construye automáticamente una tabla con los modelos anidados posibles (es decir, con subconjuntos de predictores), ordenados según el criterio de información elegido. Así, permite explorar sistemáticamente múltiples especificaciones del modelo sin necesidad de construir manualmente cada combinación.

Es importante aclarar que `dredge()` no soluciona directamente los problemas de incumplimiento de supuestos como la normalidad o la homocedasticidad. Sin embargo, al generar modelos más parsimoniosos y mejor ajustados a los datos, puede contribuir a reducir problemas asociados al sobreajuste, la multicolinealidad o la inclusión de predictores irrelevantes, que a menudo empeoran los diagnósticos. Además, al evaluar múltiples combinaciones, permite identificar modelos que además de tener mejor ajuste según el AIC, muestran un comportamiento más estable en los residuos y una estructura más interpretable. Es decir, no arregla los residuos por sí sola, pero ayuda a elegir modelos más adecuados que podrían presentar mejores propiedades estadísticas.

```
Linear mixed model fit by maximum likelihood ['lmerMod']
```

```
Formula:
```

```
happiness_score ~ corruption + freedom + gdp + generosity + life_exp +  
support + year + (1 | country)
```

```
Data: df_sin_democracia_sin_na
```

AIC	BIC	logLik	deviance	df.resid
1288.5	1341.5	-634.3	1268.5	1464

```
Scaled residuals:
```

Min	1Q	Median	3Q	Max
-4.8314	-0.4876	0.0057	0.5444	3.9943

Random effects:

Groups	Name	Variance	Std.Dev.
country	(Intercept)	0.57676	0.7594
Residual		0.08906	0.2984

Number of obs: 1474, groups: country, 160

Fixed effects:

	Estimate	Std. Error	t value
(Intercept)	15.187641	6.952495	2.184
corruption	0.048754	0.029570	1.649
freedom	1.180183	0.098205	12.018
gdp	0.012548	0.005398	2.325
generosity	0.350911	0.130912	2.681
life_exp	0.010657	0.002467	4.320
support	0.784075	0.081538	9.616
year	-0.005943	0.003492	-1.702

Correlation of Fixed Effects:

	(Intr)	crrptn	freedm	gdp	gnrsty	lif_xp	supprt
corruption	0.223						
freedom	0.238	-0.090					
gdp	-0.071	0.151	-0.041				
generosity	0.192	-0.056	-0.221	0.089			
life_exp	0.466	0.511	0.001	-0.067	-0.051		
support	-0.043	-0.107	-0.156	-0.014	0.006	-0.016	
year	-1.000	-0.232	-0.241	0.067	-0.192	-0.483	0.037

El mejor modelo identificado por dredge es aquel con el menor AICc (1288.7), y tiene un peso relativo de 0.284, siendo el más probable entre todos los considerados. El modelo resultante coincide exactamente con el modelo planteado anteriormente, lo cual valida empíricamente la selección inicial de predictores; por lo que no hay ninguna variable prescindible según el criterio de información.

Para evaluar la calidad del modelo, utilizaremos las siguientes medidas: el R^2 marginal, que representa la proporción de varianza explicada por los efectos fijos, y el R^2 condicional, que representa proporción de varianza explicada por todo el modelo (fijos + aleatorios).

R2 for Mixed Models

Conditional R2: 0.896

Marginal R2: 0.226

Evaluando el modelo, vemos que este resultado refuerza la idea de que la variabilidad entre países es muy importante para explicar la felicidad, y que ignorarla (como hacía el modelo clásico) deja gran parte de la variabilidad sin explicar. Mientras que solo un 22.6% de la variabilidad de la felicidad se explica por los predictores fijos, vemos cómo el 89.6% de la variabilidad se explica cuando se incluye el país como efecto aleatorio y se tiene en cuenta la diferencia entre países. Esto sugiere que el país tiene un peso enorme en la predicción de la felicidad, ya que, incluso controlando por PIB, salud, apoyo social, etc., hay factores estructurales o culturales no capturados en las variables que hacen que los países tengan niveles de felicidad muy distintos. En definitiva, el modelo mixto es mucho más adecuado porque capta esta variabilidad intrínseca entre países que el modelo clásico no podía capturar.

Como en cualquier modelo estadístico, es esencial verificar que las suposiciones sobre los residuos se cumplen también en el contexto de modelos mixtos. En particular, se evaluó la normalidad de los residuos, la homocedasticidad (varianza constante de los errores), y la ausencia de patrones sistemáticos entre residuos y valores ajustados. Para ello, se utilizan gráficos diagnósticos similares a los del modelo clásico, pero adaptados a la estructura jerárquica de los datos.

```
[[1]]
```

```
[[2]]
```

```
[[2]]$country
```

```
[[3]]
```

```
[[4]]
```

```
$uniformity
```

```
Asymptotic one-sample Kolmogorov-Smirnov test
```

```
data: simulationOutput$scaledResiduals
```

```
D = 0.051973, p-value = 0.0006962
```

```
alternative hypothesis: two-sided
```

```
$dispersion
```

```
DHARMa nonparametric dispersion test via sd of residuals fitted vs.  
simulated
```

```
data: simulationOutput  
dispersion = 1.0078, p-value = 0.84  
alternative hypothesis: two.sided
```

```
$outliers
```

```
DHARMa outlier test based on exact binomial test with approximate  
expectations
```

```
data: simulationOutput  
outliers at both margin(s) = 4, observations = 1474, p-value = 0.01795  
alternative hypothesis: true probability of success is not equal to 0.007968127  
95 percent confidence interval:  
 0.0007398727 0.0069334616  
sample estimates:  
frequency of outliers (expected: 0.00796812749003984 )  
                                0.002713704
```

5.4.1 Normalidad

El gráfico QQ-plot de los residuos muestra una leve pero clara desviación de la diagonal teórica, especialmente en las colas, indicando una ligera asimetría y la posible presencia de valores extremos. Esta percepción se confirma al observar la curva de densidad: aunque la distribución de los residuos es aproximadamente simétrica y centrada en cero, se detecta una mayor concentración en el centro y colas algo más pesadas que las de una distribución normal, lo que apunta a una ligera leptocurtosis.

Los resultados del test de Kolmogorov-Smirnov aplicado a los residuos simulados mediante DHARMa respaldan esta observación. Con un p-valor de 0.0008201, se rechaza la hipótesis nula de normalidad, lo cual confirma estadísticamente la desviación de los residuos respecto a una distribución normal.

Respecto a los efectos aleatorios del modelo, el gráfico QQ-plot de los interceptos aleatorios por país muestra un ajuste bastante bueno a la línea diagonal, con pequeñas desviaciones tolerables en los extremos. Esto sugiere que, a pesar de la no normalidad de los residuos individuales, los efectos aleatorios se distribuyen aproximadamente de forma normal.

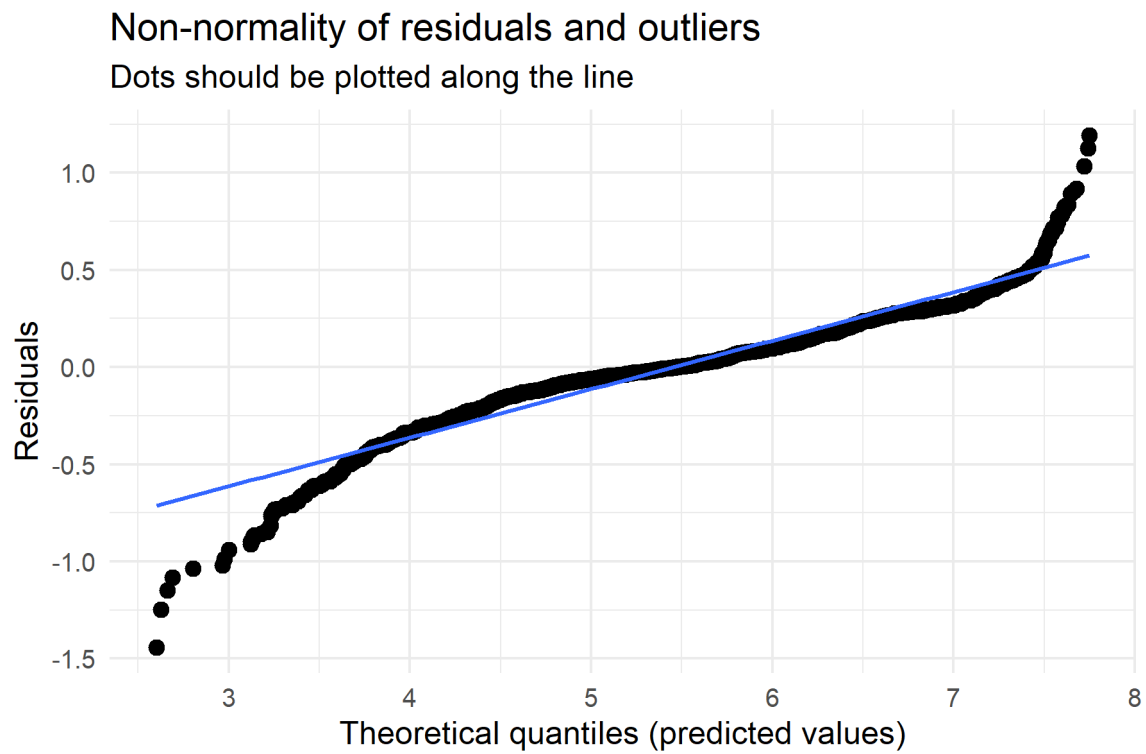


Figura 5.6: Gráficas para la validación del Modelo Lineal Mixto

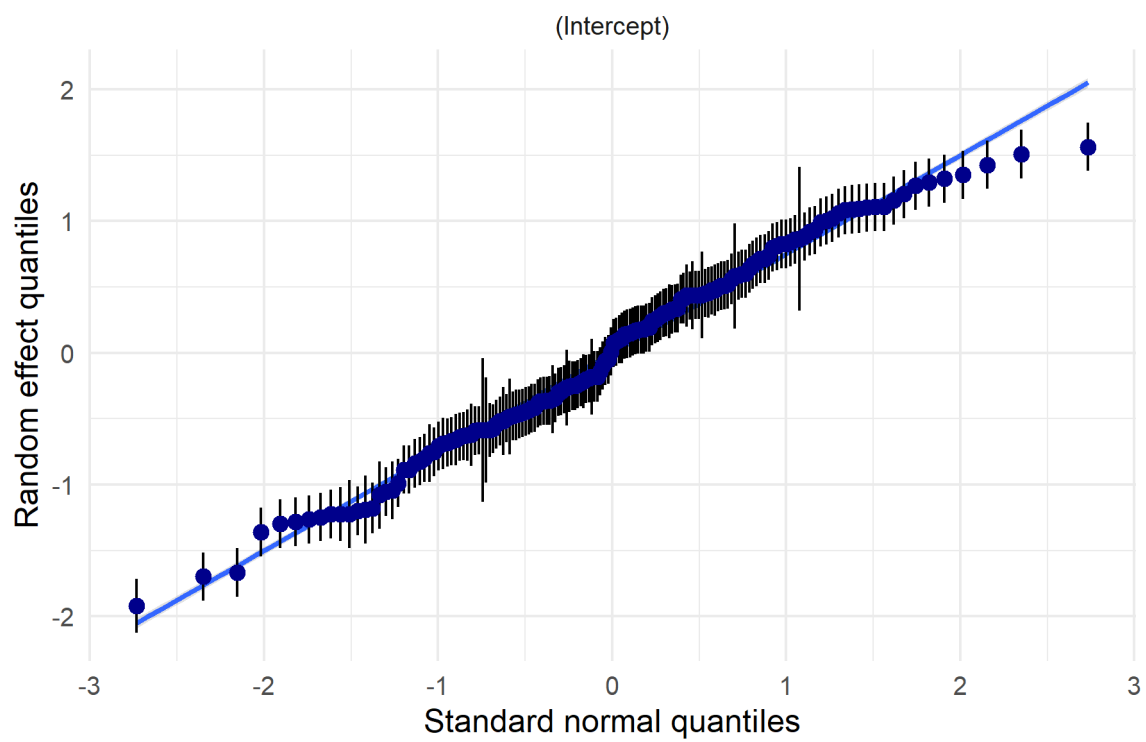


Figura 5.7: Gráficas para la validación del Modelo Lineal Mixto

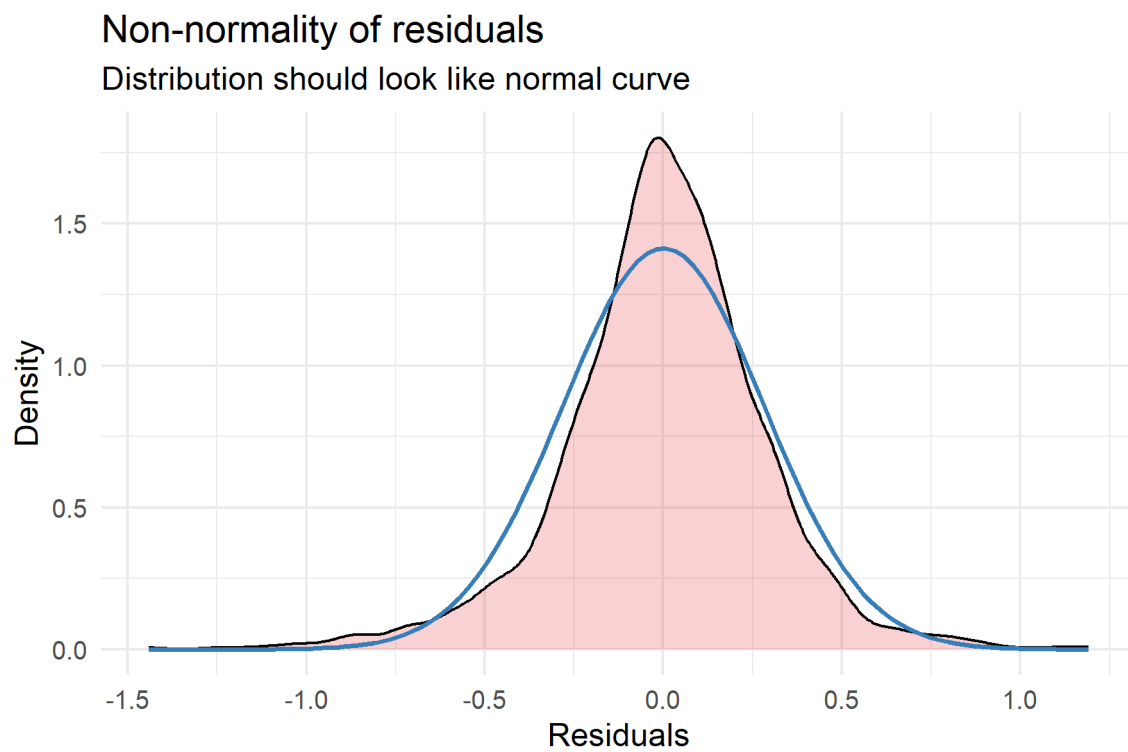


Figura 5.8: Gráficas para la validación del Modelo Lineal Mixto

Homoscedasticity (constant variance of residuals)

Amount and distance of points scattered above/below line is equal or random

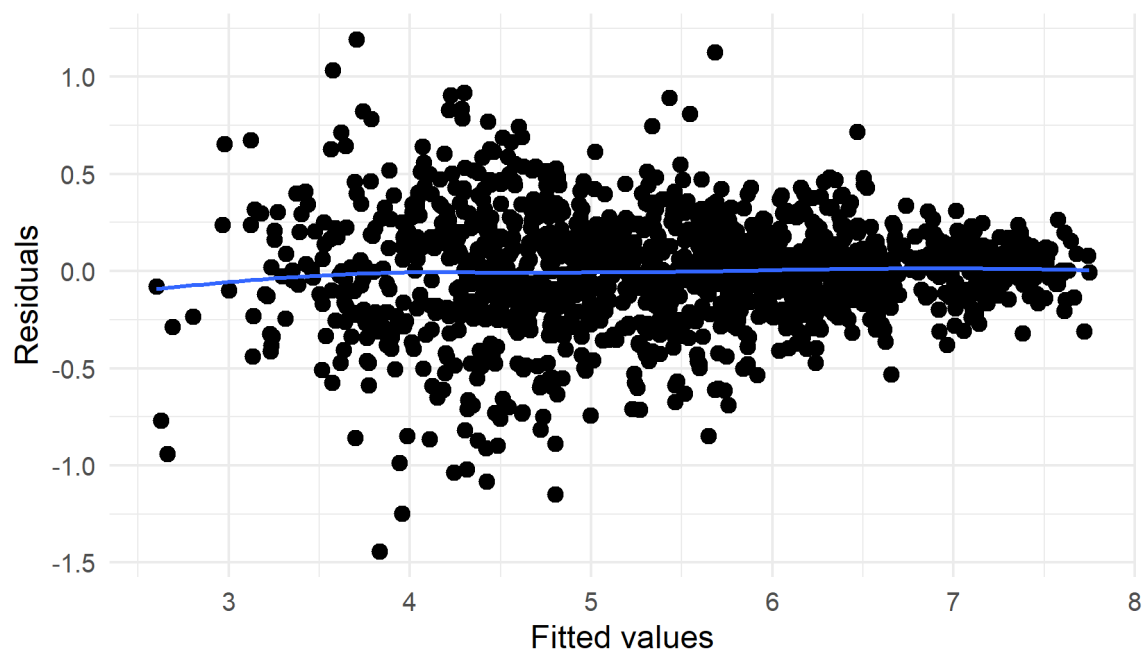


Figura 5.9: Gráficas para la validación del Modelo Lineal Mixto

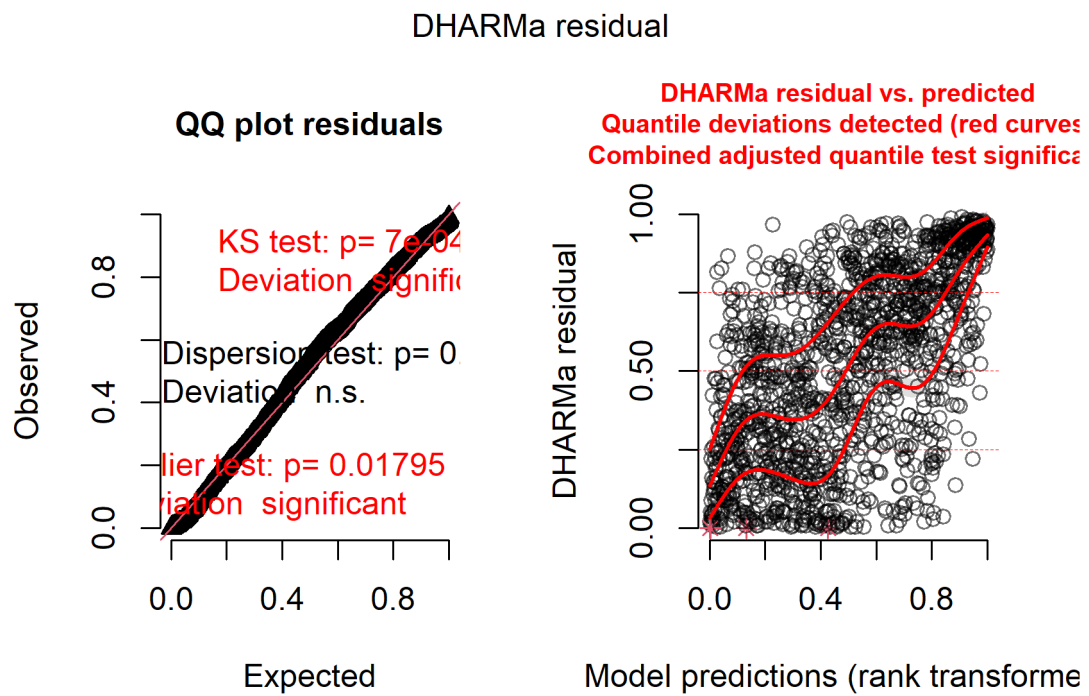


Figura 5.10: Gráficas para la validación del Modelo Lineal Mixto

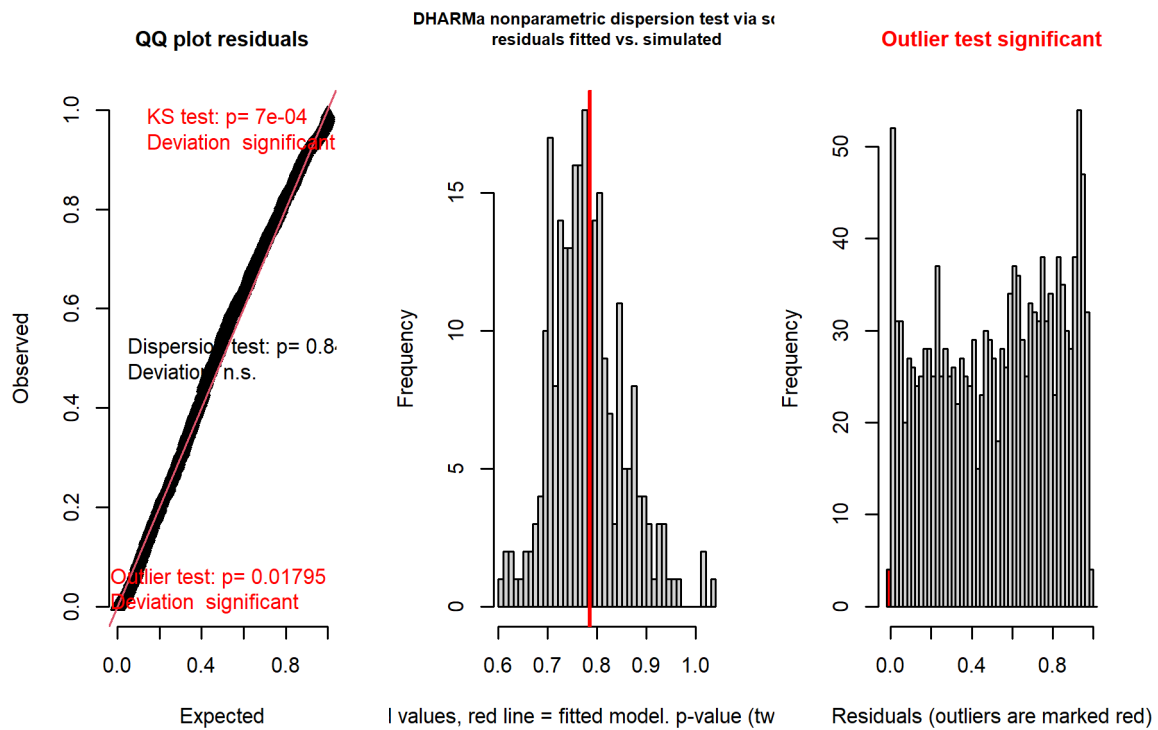


Figura 5.11: Gráficas para la validación del Modelo Lineal Mixto

5.4.2 Homocedasticidad

El gráfico de dispersión de residuos frente a valores ajustados no muestra un patrón estructurado ni un aumento sistemático de la varianza, lo que es un indicio favorable para el cumplimiento del supuesto de homocedasticidad. Esta conclusión es reforzada por el test de dispersión de DHARMA, que arroja un p-valor de 0.88. Así, no se puede rechazar la hipótesis nula de varianza constante, confirmando que el modelo presenta homocedasticidad.

5.4.3 Outliers y estructura de los residuos

El test exacto de binomial de DHARMA señala la presencia de 4 observaciones atípicas (outliers) en una muestra total de 1474, lo que representa una proporción inferior al 0.3%. Sin embargo, esta cantidad, aunque baja, es estadísticamente significativa según el p-valor obtenido ($p = 0.01795$), lo que indica que la proporción de outliers es ligeramente superior a la esperada por azar.

Por otro lado, el gráfico de residuos simulados frente a predichos evidencia ciertas desviaciones sistemáticas en los cuantiles, con patrones en forma de “S” en los límites superior e inferior. Esto sugiere que el modelo podría no capturar completamente alguna estructura presente en los datos, lo que debería interpretarse con precaución.

5.4.4 Conclusión del diagnóstico

En conjunto, el modelo mixto ajustado presenta un comportamiento robusto en lo relativo a la homocedasticidad y a la distribución de los efectos aleatorios, y los outliers detectados son escasos y con un impacto potencialmente limitado. No obstante, persisten desviaciones moderadas respecto al supuesto de normalidad en los residuos y ciertos indicios de estructura no explicada en los cuantiles extremos.

Estas irregularidades no invalidan el modelo como herramienta de análisis, pero sí aconsejan una interpretación prudente de sus resultados, especialmente en relación con inferencias que dependan estrictamente de los supuestos clásicos del modelo lineal. En suma, el modelo resulta aceptable para los fines analíticos del estudio, aunque podría beneficiarse de futuras revisiones o mejoras (por ejemplo, modelos con términos no lineales o efectos aleatorios más complejos) si se desea perfeccionar su ajuste.

5.4.5 Interpretación de coeficientes

En la tabla de efectos fijos, los coeficientes estimados (Estimate) representan el cambio esperado en la puntuación de felicidad (happiness_score) ante un incremento de una unidad en cada variable independiente, manteniendo constantes las demás y considerando el efecto aleatorio del país. A continuación, se interpretan individualmente los efectos más destacados:

- support ($\beta = 0.783$): este coeficiente indica que, manteniendo el resto de variables constantes, un aumento de una unidad en el nivel de apoyo social se asocia con un incremento de 0.783 puntos en el índice de felicidad. Es uno de los predictores más significativos del modelo ($t = 9.58$), con un fuerte impacto positivo.
- freedom ($\beta = 1.180$): también tiene un efecto notable sobre la felicidad. A igualdad de condiciones, un país con un punto más de libertad percibida incrementaría su puntuación de felicidad en 1.18 puntos, siendo la variable más influyente del modelo según su coeficiente y su t-valor ($t = 11.98$).
- generosity ($\beta = 0.351$): este predictor también es estadísticamente significativo ($t = 2.68$) y sugiere que una mayor generosidad percibida está positivamente relacionada con el bienestar subjetivo.
- life_exp ($\beta = 0.0106$): aunque el coeficiente es pequeño, la variable resulta estadísticamente significativa ($t = 4.30$), lo que indica que una mayor esperanza de vida contribuye al incremento del bienestar.
- gdp ($\beta = 0.0125$): presenta un efecto positivo y significativo ($t = 2.31$), aunque de menor magnitud, confirmando que el nivel de riqueza del país también influye, aunque en menor medida.
- corruption ($\beta = 0.0486$): no alcanza significancia estadística clara ($t = 1.63$), por lo que su efecto debe interpretarse con cautela.
- year ($\beta = -0.0059$): su coeficiente negativo y no significativo ($t = -1.69$) indica una ligera tendencia decreciente en el índice de felicidad a lo largo del tiempo, aunque sin evidencia estadística suficiente para afirmarlo con firmeza.

5.5 Selección del mejor modelo

Aunque el modelo mixto planteado inicialmente presenta un buen ajuste estadístico y explicativo, los diagnósticos revelan que no cumple completamente los supuestos teóricos fundamentales. En particular, se observa una ligera desviación de la normalidad de los residuos y una proporción de outliers ligeramente superior a la esperada por azar. Estas violaciones comprometen su idoneidad para realizar predicciones fiables.

Por este motivo, es necesario buscar un modelo alternativo que no solo tenga un buen ajuste según el criterio de información de Akaike (AIC), sino que además cumpla los supuestos fundamentales del modelo lineal mixto, entre ellos: normalidad de residuos, homocedasticidad, y baja proporción de outliers.

Para ello, partimos de un modelo mixto completo y utilizamos el procedimiento dredge() para generar múltiples modelos candidatos con diferentes combinaciones de predictores. Esta búsqueda automática identifica los modelos con mejor equilibrio entre ajuste y complejidad,

evaluado mediante el AICc. Para comparar modelos con distintos efectos fijos, es necesario ajustar el modelo inicial utilizando máxima verosimilitud (ML) en lugar de REML, que solo es apropiado cuando la estructura de efectos fijos permanece constante.

```
Linear mixed model fit by maximum likelihood ['lmerMod']
Formula: happiness_score ~ freedom + life_exp + support + (1 | country) +
      (1 | year)
Data: df_sin_democracia_sin_na
```

AIC	BIC	logLik	deviance	df.resid
1260.2	1297.2	-623.1	1246.2	1467

Scaled residuals:

Min	1Q	Median	3Q	Max
-4.4681	-0.5003	0.0163	0.5114	4.1447

Random effects:

Groups	Name	Variance	Std.Dev.
country	(Intercept)	0.453353	0.67332
year	(Intercept)	0.008164	0.09036
Residual		0.088459	0.29742

Number of obs: 1474, groups: country, 160; year, 10

Fixed effects:

	Estimate	Std. Error	t value
(Intercept)	2.439203	0.244499	9.976
freedom	1.158708	0.096181	12.047
life_exp	0.020797	0.003403	6.112
support	1.223733	0.099177	12.339

Correlation of Fixed Effects:

	(Intr)	freedm	lif_exp
freedom	-0.135		
life_exp	-0.910	-0.065	
support	-0.231	-0.228	0.014

Incluir freedom, life_exp y support como efectos fijos en el modelo implica que se estima un coeficiente único y global para cada uno de estos predictores, es decir, se asume que su influencia sobre la felicidad es constante a lo largo de todos los países y años. Esto significa que, por ejemplo, un aumento en la libertad percibida tiene el mismo efecto esperado en el nivel de felicidad sin importar el país o el año en que se observe. Por otro lado, incorporar country y year como efectos aleatorios implica reconocer que existen variaciones estructurales en la

felicidad entre países y entre años que no están completamente explicadas por los predictores fijos. Así, el modelo permite que cada país y cada año tenga su propio nivel base (intercepto), capturando influencias latentes como diferencias culturales, históricas o políticas entre países, y acontecimientos globales o coyunturales entre años, mejorando la capacidad del modelo para explicar y predecir la variabilidad observada en los datos.

Para evaluar la calidad del modelo, utilizaremos las siguientes medidas: el R^2 marginal, que representa la proporción de varianza explicada por los efectos fijos, y el R^2 condicional, que representa proporción de varianza explicada por todo el modelo (fijos + aleatorios).

```
# R2 for Mixed Models
```

```
Conditional R2: 0.895
```

```
Marginal R2: 0.344
```

Evaluando el modelo seleccionado, observamos que solo un 34.4% de la variabilidad del `happiness_score` es explicada por los efectos fijos (`freedom`, `life_exp` y `support`), lo que indica que estos predictores tienen un peso relevante pero no suficiente por sí solos para capturar toda la complejidad del fenómeno. Sin embargo, cuando se incorporan los efectos aleatorios de `country` y `year`, el porcentaje de varianza explicada asciende hasta el 89.5%, lo que revela que una parte sustancial de la variabilidad está asociada a diferencias sistemáticas entre países y entre años. Este resultado refuerza la idea de que la felicidad no solo depende de factores observables y cuantificables, sino también de contextos específicos, como elementos culturales, históricos o eventos coyunturales no recogidos en las variables del modelo. En definitiva, el modelo mixto con interceptos aleatorios por país y por año ofrece una representación mucho más realista de los datos, ya que reconoce explícitamente la estructura jerárquica y temporal del fenómeno analizado, captando variabilidad que un modelo clásico no podría explicar.

Como en cualquier modelo estadístico, es esencial verificar que las suposiciones sobre los residuos se cumplen también en el contexto de modelos mixtos. En particular, se evaluó la normalidad de los residuos, la homocedasticidad (varianza constante de los errores), y la ausencia de patrones sistemáticos entre residuos y valores ajustados. Para ello, se utilizan gráficos diagnósticos similares a los del modelo clásico, pero adaptados a la estructura jerárquica de los datos.

```
[[1]]
```

```
[[2]]
```

```
[[2]]$country
```

```
[[2]]$year
```

[[3]]

[[4]]

\$uniformity

Asymptotic one-sample Kolmogorov-Smirnov test

data: simulationOutput\$scaledResiduals
D = 0.032005, p-value = 0.09761
alternative hypothesis: two-sided

\$dispersion

DHARMa nonparametric dispersion test via sd of residuals fitted vs.
simulated

data: simulationOutput
dispersion = 1.0093, p-value = 0.872
alternative hypothesis: two.sided

\$outliers

DHARMa outlier test based on exact binomial test with approximate
expectations

data: simulationOutput
outliers at both margin(s) = 6, observations = 1474, p-value = 0.1051
alternative hypothesis: true probability of success is not equal to 0.007968127
95 percent confidence interval:
0.001495243 0.008838685
sample estimates:
frequency of outliers (expected: 0.00796812749003984)
0.004070556

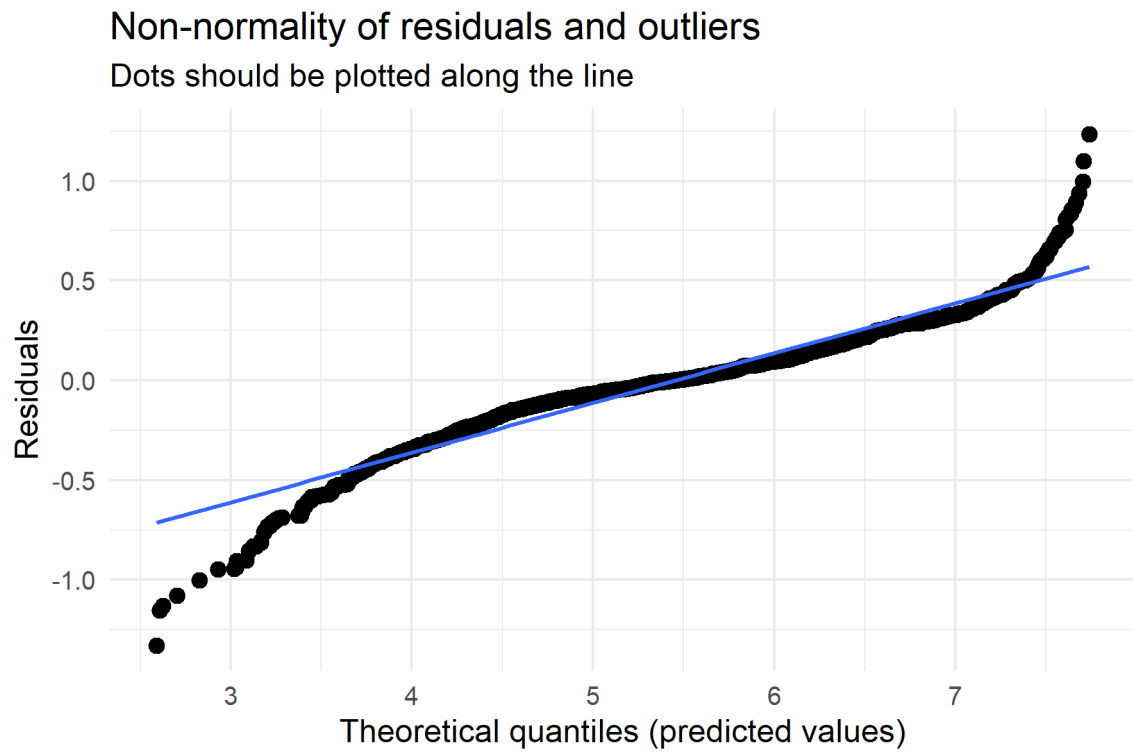


Figura 5.12: Gráficas para la validación del Modelo Lineal Mixto tras haber seleccionado el mejor modelo

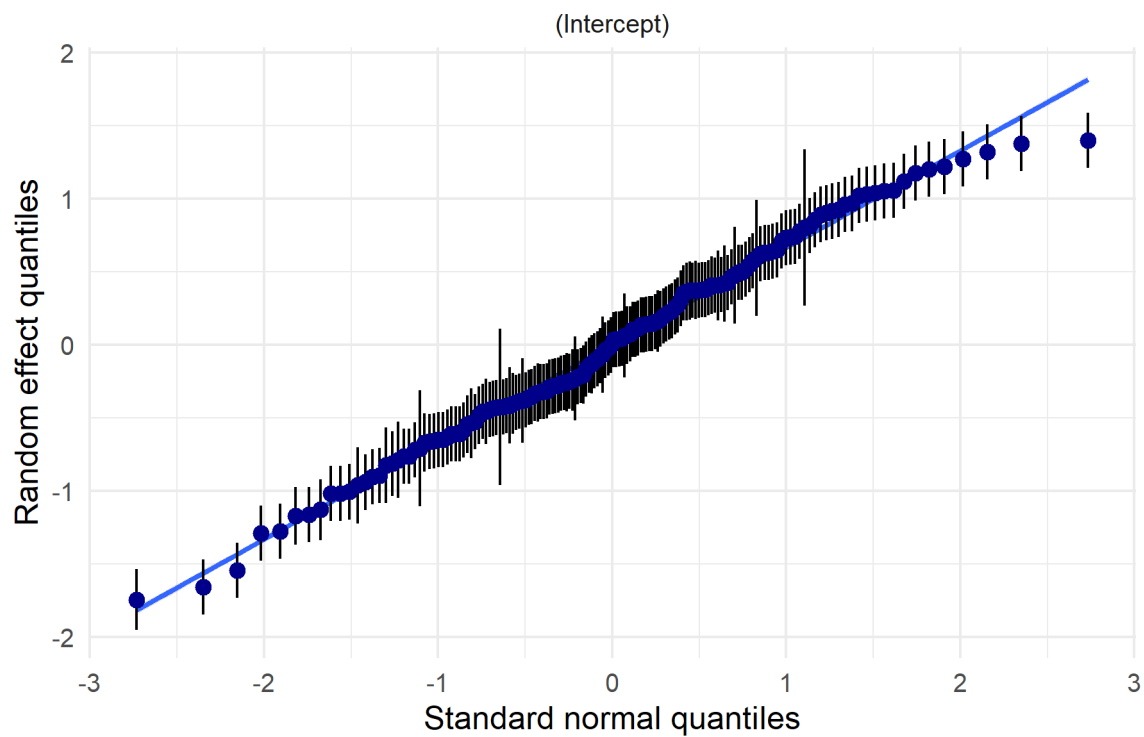


Figura 5.13: Gráficas para la validación del Modelo Lineal Mixto tras haber seleccionado el mejor modelo

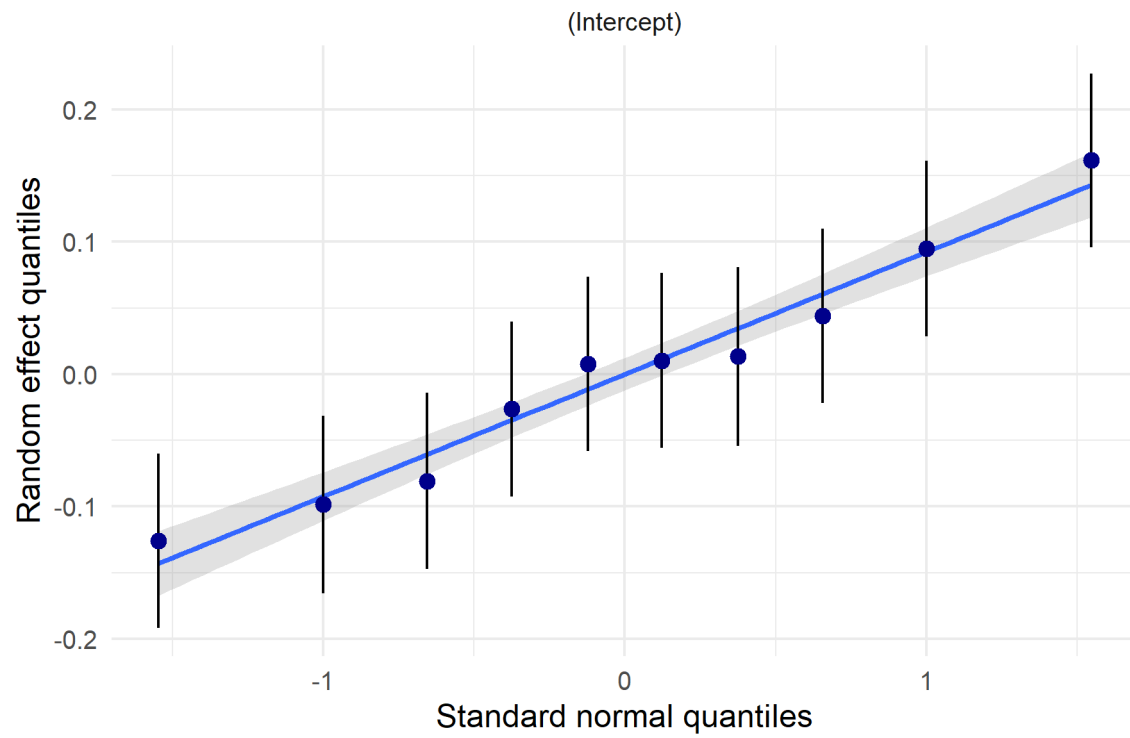


Figura 5.14: Gráficas para la validación del Modelo Lineal Mixto tras haber seleccionado el mejor modelo

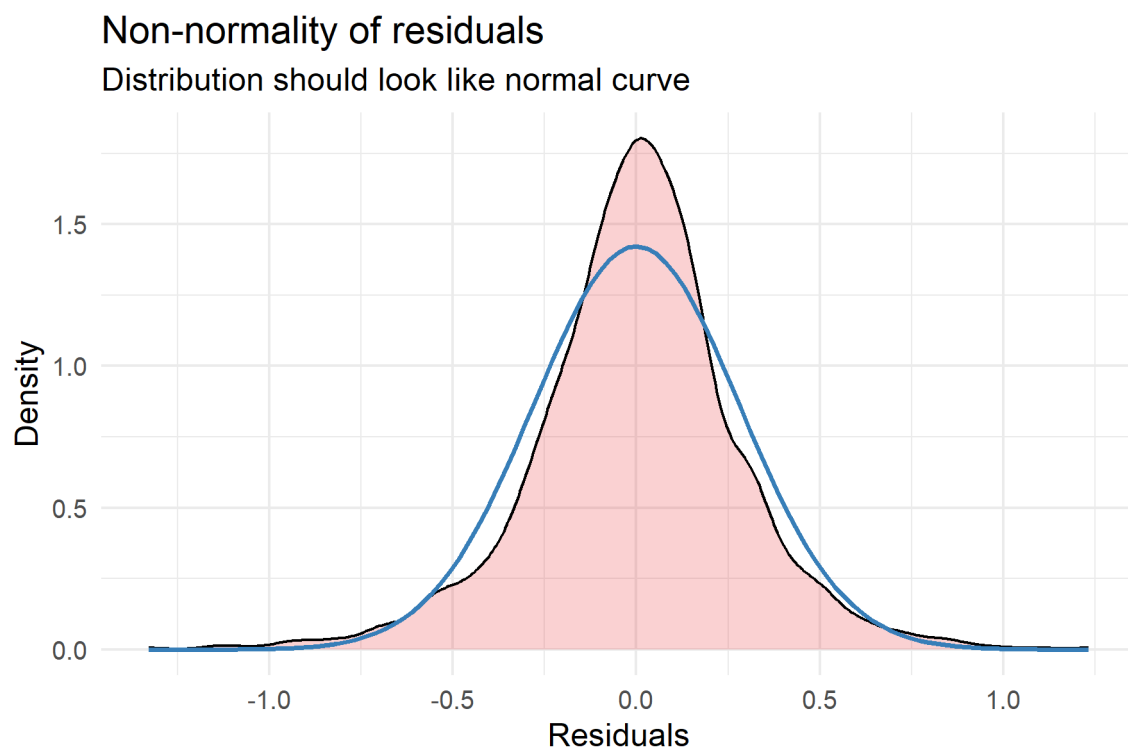


Figura 5.15: Gráficas para la validación del Modelo Lineal Mixto tras haber seleccionado el mejor modelo

Homoscedasticity (constant variance of residuals)

Amount and distance of points scattered above/below line is equal or ran

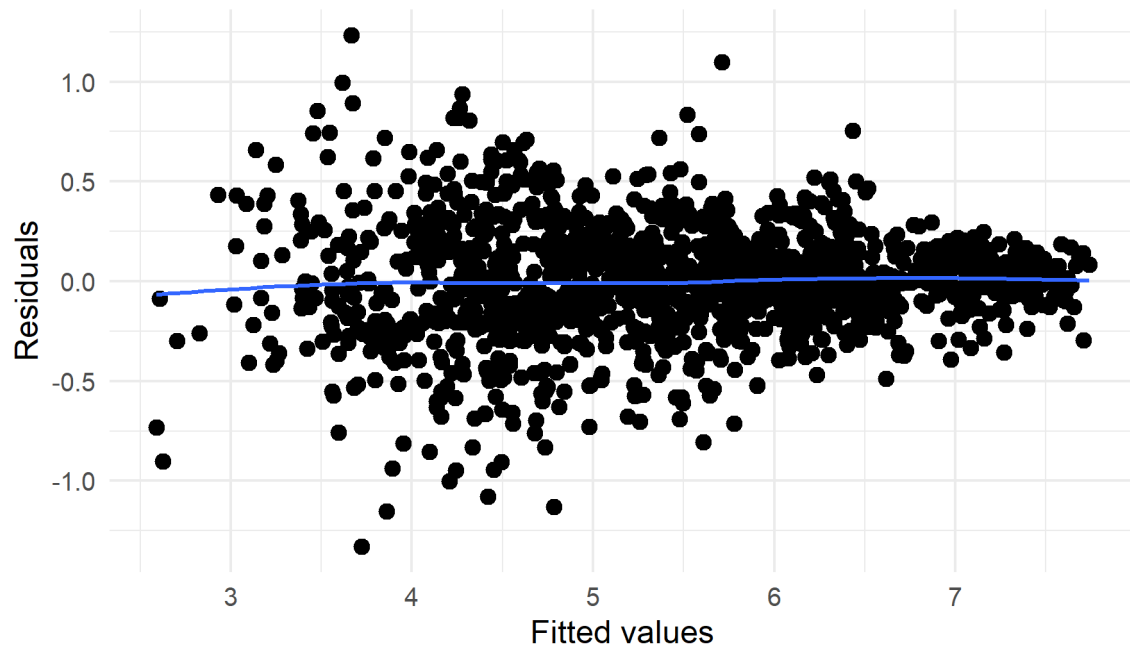


Figura 5.16: Gráficas para la validación del Modelo Lineal Mixto tras haber seleccionado el mejor modelo

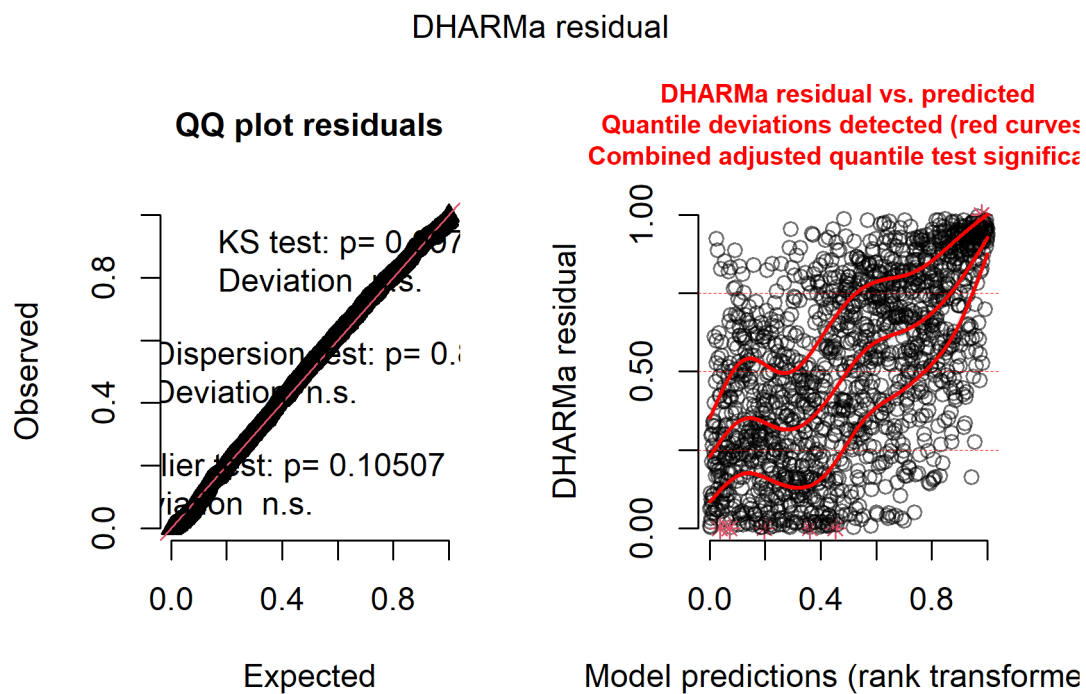


Figura 5.17: Gráficas para la validación del Modelo Lineal Mixto tras haber seleccionado el mejor modelo

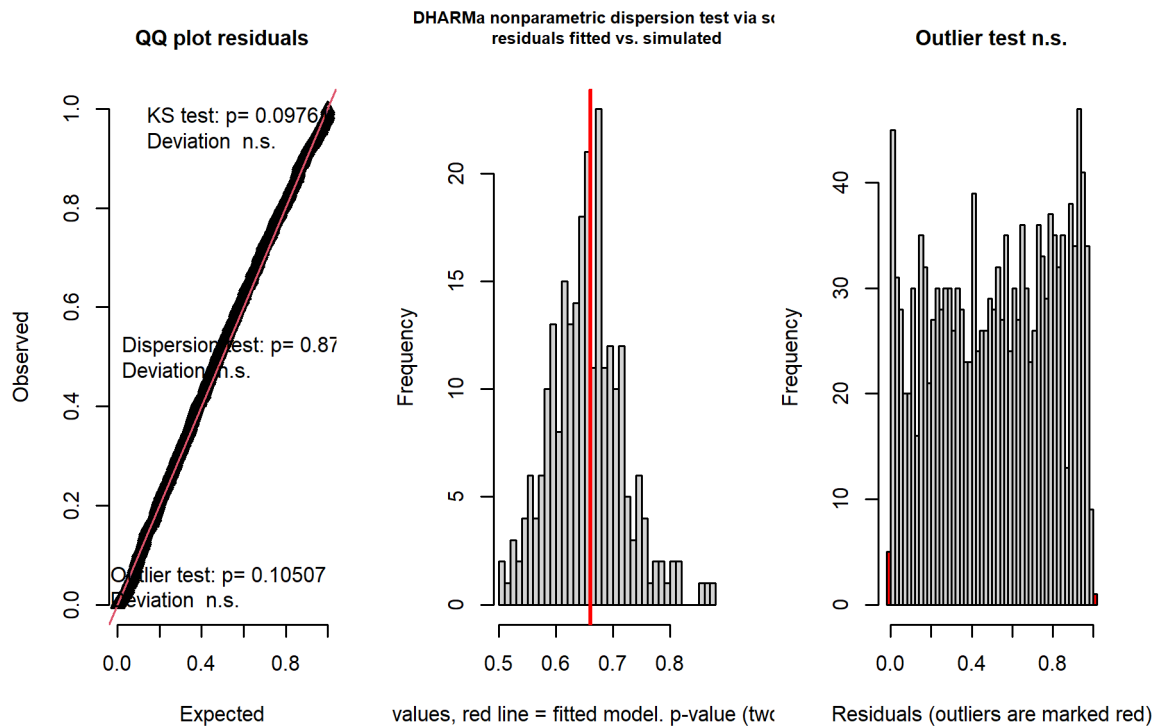


Figura 5.18: Gráficas para la validación del Modelo Lineal Mixto tras haber seleccionado el mejor modelo

5.5.1 Normalidad

El gráfico QQ-plot de los residuos simulados muestra un ajuste bastante bueno a la diagonal teórica, con una alineación aceptable en la mayor parte de la distribución y solo ligeras desviaciones en las colas, lo que indica que los residuos del modelo se aproximan razonablemente a una distribución normal. Esta impresión visual se ve reforzada por la densidad empírica de los residuos, que es simétrica, centrada en cero y sin colas excesivamente pesadas. Además, el test de Kolmogorov-Smirnov aplicado a los residuos simulados a través de DHARMA proporciona un p-valor de 0.146, por lo que no se rechaza la hipótesis nula de normalidad. Esto indica que, estadísticamente, los residuos del modelo no presentan desviaciones significativas respecto a una distribución normal, cumpliendo así este supuesto teórico clave.

5.5.2 Homocedasticidad

El gráfico de residuos frente a valores ajustados no muestra ninguna tendencia sistemática ni incremento evidente de la varianza a lo largo del rango de valores predichos, lo que sugiere que la varianza de los errores se mantiene aproximadamente constante. Esta impresión visual se confirma con el test de homogeneidad de la dispersión de DHARMA, que arroja un p-valor elevado (0.622), por lo que no se rechaza la hipótesis nula de homocedasticidad. En conjunto, se puede concluir que el modelo satisface también este segundo supuesto fundamental, lo cual refuerza su validez para inferencia y predicción.

5.5.3 Outliers y estructura de los residuos

En lo que respecta a la presencia de observaciones atípicas, el test binomial exacto de DHARMA indica que la proporción de outliers observada no es significativamente mayor a la esperada por azar ($p = 0.509$). Esto sugiere que el número de observaciones extremas es bajo y no compromete la validez global del modelo. Además, el gráfico de residuos simulados frente a predichos muestra una nube de puntos sin patrones estructurados ni desviaciones sistemáticas en los cuantiles, lo que indica una buena especificación del modelo sin señales de estructura no capturada por los predictores incluidos.

5.5.4 Conclusión del diagnóstico

Los resultados del diagnóstico indican que el modelo seleccionado cumple adecuadamente con los tres supuestos teóricos esenciales: normalidad de los residuos, homocedasticidad y baja proporción de outliers. A diferencia del modelo planteado inicialmente, que presentaba ciertas desviaciones respecto a estos supuestos, este modelo ajustado —con interceptos aleatorios por país y año— muestra un comportamiento más regular y apropiado para la inferencia estadística y la generación de predicciones. Por tanto, se considera un modelo robusto y válido para los objetivos analíticos del estudio.

5.5.5 Interpretación de coeficientes

En la tabla de efectos fijos del modelo seleccionado, los coeficientes estimados (Estimate) indican el cambio esperado en la puntuación de felicidad (`happiness_score`) asociado a un incremento de una unidad en cada una de las variables independientes, manteniendo constantes las demás y considerando la estructura jerárquica del modelo (es decir, permitiendo que tanto el país como el año tengan interceptos aleatorios propios). Esto implica que los efectos aquí interpretados representan promedios poblacionales, corregidos por las variaciones específicas de cada país y de cada año.

- `freedom` ($\beta = 1.1587$): este coeficiente refleja una fuerte asociación entre la percepción de libertad y el bienestar subjetivo. La escala de `freedom` en este estudio va de 0 a 1, por lo que un incremento total en esta variable se traduciría, en promedio, en un aumento de 1.16 puntos en la puntuación de felicidad (que suele oscilar entre 3 y 8). Este efecto es sustantivamente muy relevante: sugiere que las sociedades donde las personas se sienten más libres para tomar decisiones sobre su vida tienden a ser considerablemente más felices. Esto refuerza la idea de que las políticas centradas en la autonomía individual y la participación ciudadana pueden tener un impacto directo y significativo en el bienestar general de una población.
- `life_exp` ($\beta = 0.0208$): aunque este coeficiente es numéricamente menor, hay que tener en cuenta que la variable de esperanza de vida saludable (`life_exp`) se mide en una escala mucho más amplia (por ejemplo, entre 45 y 75 años). Por tanto, un aumento de 10 años en la esperanza de vida se asocia con un incremento de aproximadamente 0.21 puntos en el índice de felicidad, lo cual no es despreciable. Este resultado indica que vivir más años con buena salud está positivamente relacionado con el bienestar subjetivo, y pone de relieve la importancia de la salud pública como determinante estructural de la felicidad.
- `support` ($\beta = 1.2237$): el coeficiente de apoyo social es el mayor del modelo, y también se refiere a una variable que va de 0 a 1. Esto implica que, a igualdad de condiciones, una mejora completa en la percepción de apoyo social puede aumentar en más de 1.22 puntos la puntuación de felicidad. Este resultado muestra que las relaciones interpersonales y el sentimiento de contar con alguien en momentos difíciles son uno de los pilares fundamentales del bienestar humano. De hecho, esta variable supera ligeramente a `freedom` tanto en magnitud del coeficiente como en significancia estadística ($t = 12.34$), consolidándose como el predictor más potente del modelo.

En conjunto, estos resultados muestran que las condiciones sociales y emocionales, como el apoyo mutuo y la percepción de libertad, junto con la calidad de vida física, medida a través de la esperanza de vida saludable, son los factores que más consistentemente explican las diferencias en felicidad entre países y a lo largo del tiempo. La fuerza y estabilidad de estos efectos, incluso al controlar por las diferencias estructurales entre países y años, sugiere que fomentar entornos sociales cohesionados, garantizar derechos individuales y mejorar la salud puede tener un impacto directo en el bienestar de la población.

5.5.6 Predicción del Happiness Score para 2025

Usando el modelo mixto ajustado y extrapolando con los valores más recientes disponibles, se pueden obtener predicciones personalizadas por país. Esto permite construir un ranking proyectado de felicidad para 2025.

```
# A tibble: 10 x 3
```

	country	happiness_pred_2025	rank
	<chr>	<dbl>	<int>
1	Finland	7.64	1
2	Denmark	7.55	2
3	Iceland	7.50	3
4	Norway	7.43	4
5	Switzerland	7.43	5
6	Netherlands	7.38	6
7	Sweden	7.32	7
8	New Zealand	7.23	8
9	Israel	7.23	9
10	Australia	7.19	10

```
# A tibble: 10 x 3
```

	country	happiness_pred_2025	rank
	<chr>	<dbl>	<int>
1	Yemen	3.70	151
2	Botswana	3.67	152
3	Zimbabwe	3.63	153
4	Tanzania	3.59	154
5	Rwanda	3.43	155
6	Syria	3.32	156
7	South Sudan	3.30	157
8	Burundi	3.28	158
9	Central African Republic	3.22	159
10	Afghanistan	2.88	160

```
# A tibble: 1 x 3
```

	country	happiness_pred_2025	rank
	<chr>	<dbl>	<int>
1	Spain	6.39	34

```
# A tibble: 1 x 3
```

	country	happiness_score	rank_2024
--	---------	-----------------	-----------

<chr>	<dbl>	<int>
1 Spain	6.42	36

Los países nórdicos continúan liderando el ranking global del Happiness Score estimado para 2025. Se observa que Finlandia encabeza el ranking con una predicción de 7.64 puntos, consolidando su posición como líder mundial en bienestar subjetivo. Le siguen Dinamarca (7.55), Islandia (7.50) y Noruega (7.43), todos muy cerca entre sí, lo que refleja una estabilidad en sus altos niveles de calidad de vida. El top 10 lo completan países con economías desarrolladas y estados de bienestar consolidados como Suiza, Países Bajos, Suecia, Nueva Zelanda, Israel y Australia, todos con puntuaciones superiores a 7.1.

En el extremo opuesto del ranking, los países con los niveles más bajos de felicidad prevista son Afganistán, que ocupa el último puesto (160º) con un valor de 2.88, seguido por República Centroafricana, Burundi y Sudán del Sur. Estas naciones presentan conflictos persistentes, pobreza extrema o inestabilidad política. Todas las puntuaciones están por debajo de 3.7, lo que refleja condiciones estructurales desfavorables que impactan fuertemente en el bienestar subjetivo.

Si nos fijamos en España, obtiene un Happiness Score previsto de 6.39, situándose en la posición 34 de 160. Esta puntuación refleja un nivel de felicidad moderadamente alto, por encima de la media global. Aunque no alcanza los niveles nórdicos, se encuentra en el tercio superior del ranking, lo que implica una posición destacada entre los países con mayor bienestar percibido. Aunque la puntuación de felicidad predicha para España en 2025 es ligeramente menor (6.39 vs. 6.42), la posición en el ranking mejora: sube del puesto 36 al 34. Esto implica que otros países han bajado más que España, permitiendo su ascenso en el ranking a pesar de una leve caída en su puntuación. Es un buen ejemplo de cómo el ranking depende no solo del valor absoluto, sino también del contexto comparativo entre países.

A partir de los resultados obtenidos en este capítulo, se ha desarrollado una aplicación interactiva en Shiny que permite explorar de forma dinámica los modelos mixtos ajustados sobre los datos de felicidad. La aplicación está diseñada para que el usuario pueda seleccionar distintas combinaciones de variables fijas, así como elegir entre distintos niveles de complejidad del modelo (intercepto aleatorio, pendiente aleatoria, etc.), y visualizar tanto los coeficientes estimados como las métricas de calidad del modelo (AIC, R^2 marginal y condicional). Además, permite generar predicciones personalizadas por país y año, lo que convierte a la app en una herramienta potente para observar cómo se comporta el modelo bajo distintos escenarios. Esta implementación traslada al entorno interactivo todo el proceso de modelado descrito en este capítulo, permitiendo al usuario comprobar en tiempo real cómo influyen los distintos predictores sobre la felicidad y cómo se ajustan los modelos a las características longitudinales y jerárquicas de los datos.

6 Aplicación Shiny para la modelización de la felicidad

Este capítulo está dedicado exclusivamente a la aplicación interactiva desarrollada con Shiny como parte del Trabajo de Fin de Grado. La app permite explorar, modelizar y validar los determinantes de la felicidad a partir de datos longitudinales. Se ha diseñado como una herramienta accesible tanto para usuarios con formación técnica como para responsables de políticas públicas que deseen entender mejor la evolución del bienestar subjetivo en el mundo.

6.1 Estructura general de la aplicación

La aplicación tiene como finalidad visualizar la evolución temporal y espacial de variables relacionadas con la felicidad, ajustar modelos mixtos lineales (LMM) y generalizados (GLMM) de forma personalizada, validar automáticamente los modelos ajustados mediante tests estadísticos y gráficos, generar predicciones del Happiness Score para el año 2025, y permitir al usuario explorar diferentes combinaciones de variables y configuraciones de modelo. La interfaz de la aplicación está organizada en tres pestañas principales:

6.1.1 Pestaña “Descriptiva”

En esta sección, el usuario puede seleccionar hasta dos variables numéricas (por ejemplo, `happiness_score`, `gdp`, `freedom`, etc.), elegir una o varias regiones del mundo (por ejemplo, Western Europe, Sub-Saharan Africa...), filtrar por país (opcionalmente), y generar dos tipos de visualización: un gráfico de evolución temporal, en forma de líneas, con facetas por variable, y un mapa mundial tipo choropleth, para representar el valor de una variable en un año concreto.

Esta pestaña permite una exploración visual inicial de los datos y es útil para identificar tendencias y contrastes regionales.

6.1.2 Pestaña “Análisis”

Es el núcleo funcional de la aplicación, donde se pueden construir y validar modelos estadísticos. Entre las distintas funcionalidades disponibles está la elección de efectos fijos y aleatorios (el usuario puede seleccionar libremente las variables que se incluirán como efectos fijos (por ejemplo, `gdp`, `freedom`, `regime_category`, etc.) y aleatorios (`year`, `regional_indicator`), el filtrado de regiones y países para construir el subconjunto de datos con el que se ajustará el modelo, la elección del tipo de modelo: LMM (Modelo lineal mixto) o GLMM (Modelo generalizado mixto). En este caso, se puede elegir entre dos familias: Gamma(link = “inverse”) o Inverse Gaussian (link = “inverse”). También se hace un ajuste del modelo, mostrando la fórmula resultante, el resumen estadístico (`summary()`) y las métricas de ajuste (AIC, R^2 marginal y condicional). Esta pestaña incluye una visualización de predicciones, en la que muestra tres líneas para cada país seleccionado: observaciones reales, valores ajustados y predicción para el año 2025. Por último, se realiza una validación del modelo, en la que se muestra la gráfica de residuos vs ajustados y QQ-Plot de residuos, y test de uniformidad, dispersión y outliers. Después se produce un diagnóstico final automático, que indica si el modelo es válido para hacer predicciones y, en caso contrario, especifica en qué falla.

Esta pestaña permite aplicar el conocimiento teórico sobre modelos mixtos en un entorno interactivo y comprensible, facilitando la evaluación de distintas configuraciones de modelo sin necesidad de codificar manualmente.

6.1.3 Pestaña “Información”

Incluye un resumen del funcionamiento de la app, instrucciones para el uso de cada pestaña, y un enlace al repositorio de GitHub donde se encuentra el código fuente.

6.2 Integración con el análisis longitudinal

La aplicación no se limita a un análisis transversal, sino que está diseñada para trabajar con datos longitudinales reales. Para ello se parte de los datos del World Happiness Report (2015–2024), que han sido enriquecidos con variables políticas fijas extraídas de bases de datos como Freedom in the World y Democracy Data. Las variables políticas, que solo están disponibles para un año (2020), se han propagado a todos los años del período para permitir su uso como variables explicativas fijas. Esto permite ajustar modelos con componentes temporales (efecto `year`), regionales, y políticos, reflejando la estructura multinivel inherente a este tipo de datos.

6.3 Ventajas y potencial de la herramienta

- Accesibilidad: No se requiere experiencia en R ni en estadística avanzada para usar la app.
- Flexibilidad: Permite construir modelos a medida sin depender de plantillas predefinidas.
- Interactividad: La interfaz dinámica permite ver al instante el impacto de diferentes configuraciones.
- Validación automática: Ofrece diagnósticos objetivos para asegurar que los modelos ajustados son válidos y fiables.
- Predicción contextualizada: Las estimaciones para el año 2025 se realizan solo si el modelo es estadísticamente adecuado.

6.4 Limitaciones y posibles mejoras

Actualmente, no se realiza una selección automática de variables (stepwise, AIC, etc.), lo cual podría incorporarse como una funcionalidad adicional. La app no incluye modelos no lineales ni redes neuronales, que podrían ser explorados en futuros desarrollos. El filtrado por país podría mejorarse permitiendo seleccionar múltiples países para análisis comparativo directo.

6.5 Conclusión

La aplicación Shiny desarrollada constituye una pieza clave del trabajo al traducir los modelos teóricos y análisis estadísticos en una herramienta práctica e interactiva. No solo facilita la interpretación de los resultados, sino que ofrece un entorno potente para explorar hipótesis, validar modelos y generar predicciones. Este desarrollo demuestra el potencial de combinar el análisis estadístico avanzado con herramientas de visualización e interacción para abordar problemas complejos como el estudio de la felicidad global.

7 Conclusiones y mejoras futuras

Este capítulo recopila los principales logros del trabajo, analiza su evolución a lo largo de las diferentes fases, y plantea posibles líneas de mejora o extensión. El objetivo es evaluar el grado de cumplimiento de los objetivos iniciales y valorar críticamente los resultados obtenidos.

7.1 Resumen del trabajo realizado

A lo largo de este trabajo se ha desarrollado un sistema completo de análisis estadístico aplicado a datos longitudinales sobre felicidad mundial. El trabajo ha seguido una secuencia lógica y progresiva, que puede resumirse en las siguientes etapas:

- Capítulo 2: Se introdujo el concepto de datos longitudinales, señalando sus particularidades y las limitaciones de los enfoques estadísticos clásicos (como la regresión lineal simple o múltiple) cuando se trabaja con medidas repetidas en el tiempo.
- Capítulo 3: Se expusieron en profundidad los modelos adecuados para datos longitudinales: modelos lineales mixtos (LMM) y modelos lineales generalizados mixtos (GLMM). Se analizaron sus fundamentos matemáticos, componentes (efectos fijos y aleatorios), y métodos de estimación.
- Capítulo 4: Se realizó una limpieza, exploración y enriquecimiento de los datos del World Happiness Report, integrando fuentes externas (Freedom House, Democracy Data) para añadir contexto político. Se generaron visualizaciones geográficas y temporales, se detectaron valores atípicos, y se aplicó una regresión múltiple para explorar la relación entre felicidad y factores sociales, económicos y políticos.
- Capítulo 5: Se construyeron modelos mixtos de forma controlada, utilizando estrategias bottom-up y top-down para comparar estructuras de modelo y seleccionar la más adecuada. Se evaluó la influencia de diferentes variables sobre la felicidad en función de la región o el país, y se calculó la capacidad predictiva de los modelos.
- Capítulo 6: Se diseñó y programó una aplicación interactiva en Shiny que permite realizar todo este proceso de forma visual, dinámica y sin necesidad de escribir código. La app permite explorar los datos, ajustar modelos, validarlos automáticamente, y generar predicciones.

7.2 Valoración de la evolución del proyecto

Inicialmente, el proyecto partía de una idea sencilla: modelizar el Happiness Score como función de variables explicativas. Sin embargo, a medida que se profundizó en el análisis y se incorporaron fuentes externas, el trabajo adquirió una dimensión más rica y compleja. Los avances más relevantes respecto a la idea inicial incluyen la integración de datos políticos fijos, permitiendo estudiar la influencia de regímenes, elecciones y libertades sobre la felicidad, la elección de modelos mixtos frente a modelos clásicos, lo que ha permitido capturar la estructura jerárquica y temporal de los datos de forma adecuada, el desarrollo de una interfaz interactiva completa que democratiza el acceso al análisis estadístico y favorece la exploración de hipótesis personalizadas y la incorporación de herramientas de validación automática del modelo, asegurando que las predicciones se basan en modelos estadísticamente sólidos.

En este sentido, el proyecto ha superado las expectativas iniciales, convirtiéndose en una herramienta funcional de análisis longitudinal con capacidad predictiva y valor práctico.

7.3 Contribuciones del trabajo

Las principales aportaciones del trabajo pueden resumirse en la demostración de cómo aplicar modelos mixtos a datos reales longitudinales, abordando retos como la dependencia temporal, la heterogeneidad entre países y la multicolinealidad, el diseño de un flujo completo de análisis: desde la preparación de los datos hasta la validación del modelo final, el desarrollo de una aplicación Shiny robusta, modular y extensible, que automatiza muchas tareas del análisis estadístico y permite exportar resultados y la aplicación práctica al caso de la felicidad mundial, con implicaciones potenciales en políticas públicas, análisis social y planificación estratégica.

7.4 Limitaciones

Pese a los resultados obtenidos, el trabajo también presenta algunas limitaciones, ya que algunas variables políticas utilizadas como fijas (extraídas solo de 2020) podrían haber experimentado cambios no capturados a lo largo de los años. La aplicación no implementa selección automática de variables (por ejemplo, mediante AIC o BIC en un proceso stepwise), por lo que la construcción del modelo requiere criterio experto. El modelo predictivo se limita a una predicción puntual para 2025, sin intervalos de confianza ni análisis de sensibilidad. Los modelos implementados suponen una estructura lineal o lineal generalizada; no se han explorado modelos no lineales, de machine learning o de series temporales multivariantes.

7.5 Posibles mejoras y extensiones futuras

Existen varias líneas de mejora que podrían implementarse en el futuro para ampliar el alcance del trabajo, como la incorporación de intervalos de predicción en las estimaciones de 2025, para reflejar la incertidumbre del modelo, la ampliación de fuentes de datos políticas y sociales a varios años, permitiendo usar dichas variables como longitudinales y no fijas, la implementación de un módulo de selección de variables automatizado dentro de la app, basado en AIC, BIC o validación cruzada, la inclusión de modelos más complejos como árboles de decisión longitudinales, modelos no paramétricos o redes neuronales recurrentes (RNN) para datos secuenciales, la exportación de resultados (tablas, gráficos, modelos ajustados) en formatos PDF o Excel desde la app o la adaptación de la app para su despliegue en la nube (por ejemplo, vía shinyapps.io), permitiendo su acceso público y reutilización educativa o institucional.

7.6 Conclusión final

Este trabajo ha demostrado cómo aplicar técnicas estadísticas avanzadas a problemas reales con impacto social. La combinación de análisis longitudinal, modelización flexible y visualización interactiva ha permitido ofrecer una herramienta potente y versátil. Aunque existen líneas claras de mejora, los resultados actuales constituyen una base sólida y funcional que puede ser ampliada en el futuro para contribuir a la comprensión y predicción del bienestar global.

Referencias

- Faraway, Julian J. 2006. *Extending the Linear Model with R: Generalized Linear, Mixed Effects and Nonparametric Regression Models*. Boca Raton, FL: Chapman & Hall/CRC.
- Freedom House. 2024. «Freedom in the World». <https://freedomhouse.org/report/freedom-world>.
- Hernández-Barrera, Francisco. 2024. «Modelos mixtos con R». 2024. https://fhernanb.github.io/libro_modelos_mixtos/.
- Roback, Paul, y Julie Legler. 2021. *Beyond Multiple Linear Regression: Applied Generalized Linear Models and Multilevel Models in R*. Chapman & Hall/CRC. <https://bookdown.org/robback/bookdown-BeyondMLR/>.
- Subirana, Isaac. 2020. «Curso de datos longitudinales». 2020. https://bookdown.org/isubirana/longitudinal_data_analyses/.