



**FACULTAD DE ESTUDIOS ESTADÍSTICOS**

**GRADO EN ESTADISTICA APLICADA**

**Curso 2024/2025**

---

**Trabajo de Fin de Grado**

***TÍTULO: La Masculinidad en reconstrucción:  
Análisis de la Construcción social de la  
Masculinidad en los jóvenes de España***

***Alumna: Lucía Paíno López***

***Tutora: Silvia Pineda San Juan***

Junio de 2025



UNIVERSIDAD COMPLUTENSE  
MADRID



## Contenido

Figuras.....	6
Tablas .....	7
1. Abstract .....	8
2. Introducción y objetivos .....	9
3. Metodología .....	11
3.1. Algoritmo MissForest .....	11
3.2. Matriz correlación de Pearson .....	12
3.3. Análisis Clúster .....	12
3.3.1. Reducción dimensionalidad: Análisis Componentes Principales .....	13
3.4. Regresión logística multinomial.....	13
3.4.1. Métodos de selección de variables .....	14
3.4.2. Interpretación de los parámetros .....	14
3.5. Árbol de decisión.....	14
3.6. Monte Carlo Cross-Validation (MCCV).....	15
4. Depuración de datos.....	17
4.1. Lectura de datos .....	17
4.2. Validación NS/NC.....	20
4.3. Tratamiento valores NA.....	21
5. Análisis Descriptivo.....	24
5.1. Perfil sociodemográfico .....	24
5.2. Bienestar y satisfacción personal .....	25
5.3. Presiones sociales y conductas de riesgo.....	27
5.4. Entorno social.....	29
5.5. Percepciones de masculinidad.....	30
5.6. Otras relaciones.....	31
6. Análisis Clúster .....	33
6.1. Caracterización de los clústeres.....	34
6.2. Perfilado de los clústeres .....	35
7. Análisis estadístico .....	43
7.1. Monte Carlo Cross-Validation.....	43
7.2. Modelo regresión logística multinomial.....	46
7.3. Árbol de decisión.....	49
7.3.1. Interpretación del árbol de decisión .....	52
7.3.2. Importancia de variables.....	53
8. ¿Es la masculinidad frágil perjudicial para los hombres?.....	55

9. ¿Hay diferencias entre hombres y mujeres en cuanto a la percepción de la masculinidad?.....	56
10. ¿Cómo perciben los jóvenes la masculinidad? .....	57
Conclusión .....	58
Bibliografía .....	59

## Figuras

Figura 1: Valores NA por variables .....	21
Figura 2: Valores NA por observaciones .....	21
Figura 3: Distribución por edad .....	24
Figura 4: Distribución por género .....	24
Figura 5: Distribución por clase social.....	24
Figura 6: Distribución por religiosidad .....	25
Figura 7: Distribución por ideología.....	25
Figura 8: Distribución por orientación sexual .....	25
Figura 9: Distribución por estado de felicidad .....	25
Figura 10: Distribución por grado de sentirse una/o misma/o .....	26
Figura 11: Distribución por salud mental .....	26
Figura 12: Distribución por amistades .....	26
Figura 13: Distribución por presiones.....	27
Figura 14: Distribución por índice de riesgo .....	28
Figura 15: Distribución por frecuencia de burla.....	28
Figura 16: Distribución por frecuencia de avergonzar .....	28
Figura 17: Distribución por frecuencia de violencia .....	29
Figura 18: Distribución por feminismo.....	29
Figura 19: Distribución por transmisión roles de género .....	29
Figura 20: Distribución por percepción masculinidad.....	30
Figura 21: Matriz correlaciones Pearson de variables MASC.....	31
Figura 22: Distribución por percepción de masculinidad y género .....	32
Figura 23: Método del codo .....	33
Figura 24: Visualización clústeres .....	34
Figura 25: Media de respuestas de las variables MASC por clúster .....	34
Figura 26: Distribución por clúster .....	35
Figura 27: Distribución clústeres por género.....	36
Figura 28: Distribución clase social por clúster.....	37
Figura 29: Distribución felicidad por clúster.....	39
Figura 30: Distribución de sentirse una/o misma/o por clúster .....	39
Figura 31: Distribución de salud mental por clúster.....	40
Figura 32: Distribución salud mental (hombres) por clúster .....	41
Figura 33: Nivel medio de presión social percibida según clúster .....	41
Figura 34: Mapa de flujo de Monte Carlo Cross-Validation .....	43
Figura 35: Comparación AUC entre modelos con boxplot .....	44
Figura 36: Árbol maximal.....	50
Figura 37: Error validacion cruzada según tamaño del árbol.....	50
Figura 38: Valor AUC por número hoja .....	51
Figura 39: Árbol podado .....	52
Figura 40: Importancia relativa variables árbol.....	54

## Tablas

Tabla 1: Clasificación variables 1 .....	17
Tabla 2: Clasificación variables 2 .....	17
Tabla 3: Clasificación variables 3 .....	18
Tabla 4: Clasificación variables 4 .....	18
Tabla 5: Clasificación variables 5 .....	19
Tabla 6: Clasificación variables 6 .....	19
Tabla 7: Clasificación variables 7 .....	19
Tabla 8: Clasificación variables 8 .....	20
Tabla 9: Errores imputación .....	22
Tabla 10: Errores de imputación normalizados .....	23
Tabla 11: Media de respuestas por presiones .....	27
Tabla 12: Prueba independencia chi-cuadrado clúster y género .....	36
Tabla 13: Prueba independencia chi-cuadrado clase social y clúster .....	37
Tabla 14: Prueba independencia chi-cuadrado orientación y clúster .....	37
Tabla 15: Test Kruskal-Wallis de edad y clúster .....	38
Tabla 16: Prueba independencia chi-cuadrado salud mental mujeres y clúster .....	40
Tabla 17: Prueba independencia chi-cuadrado salud mental hombres y clúster .....	40
Tabla 18: Variables incluidas por modelo de regresión .....	45
Tabla 19: OR del modelo de regresión logística multinomial con categoría “tóxica” como referencia .....	47

## 1. Abstract

La masculinidad, lejos de ser un rasgo biológico, es un constructo social que impone roles y expectativas en la sociedad que afectan a mujeres, hombres y personas que se salen de este binarismo. Este trabajo, analiza la percepción de la masculinidad en los jóvenes de España, teniendo como objetivo identificar perfiles ideológicos o comportamentales asociados a la percepción del concepto masculinidad. A través de un análisis clúster, se segmenta la percepción de masculinidad en tres perfiles, tóxica, neutra y sana. Posteriormente, mediante modelos explicativos como modelos de regresión logística multinomial y árboles de clasificación, se exploran las relaciones entre los tres perfiles de masculinidad y diversas variables sociales y conductuales.

---

Masculinity is not a biological trait. It is a social construct that imposes roles and expectations on society that affect women, men and people of other gender. This paper analyses the perception of masculinity among young people in Spain, with the aim of identifying ideological or behavioural profiles associated with the perception of the concept of masculinity. Through a cluster analysis, the perception of masculinity is segmented into three profiles: toxic, neutral and healthy. Subsequently, using explanatory models such as multinomial logistic regression models and classification trees, the relationships between the three profiles of masculinity and various social and behavioural variables are explored.

## 2. Introducción y objetivos

Al contrario de lo que mucha gente piensa, la sociedad machista en la que aún vivimos no solo perjudica a las mujeres, los hombres también se ven atrapados en una serie de constructos sociales que les imponen cargas emocionales y sociales, las cuales pueden resultar perjudiciales en su desarrollo.

La masculinidad no es un rasgo biológico, sino un constructo social que se desarrolla en un contexto cultural e histórico determinado. Según Raewyn Connell (1995) la masculinidad es una posición dentro de un sistema de relaciones de género, las prácticas asociadas a esta posición, y los efectos que estas producen. Todo ello afecta a la percepción sobre feminidad y masculinidad que tiene la sociedad. Existen diversas creencias sociales tradicionales en torno a los roles de género, por ejemplo, que los hombres deben ser personas duras, fuertes y valientes, o que las mujeres son débiles, sentimentales y sumisas. Estas creencias son impuestas desde edades tempranas, lo que lleva a que muchos niños, entendidos aquí como varones, crezcan dando por hecho ciertos privilegios o poderes, generando una desigualdad desde la niñez con respecto a las mujeres y a aquellos hombres, o personas de otro género, que no cumplen con el estereotipo impuesto.

En todas las sociedades, existe un concepto de masculinidad predominante, denominada como masculinidad hegemónica. Este concepto surgió a finales de la década de 1980. Connell lo describió como una ideología que asocia ciertos privilegios a algunos hombres, marcando ciertas actitudes como dominantes y por tanto determinando otras formas de masculinidad como inadecuadas o inferiores.

En consecuencia, estas actitudes generan efectos negativos en la sociedad independientemente del género. En particular, según datos del Instituto Nacional de Estadística (INE), la tasa de mortalidad por lesiones debidas a accidentes de tráfico fue igual a 6,009 por cada 100.000 hombres y 1,634 por cada 100.000 mujeres en 2023, lo que indica una clara diferencia por género y una mayor mortalidad en hombres, lo que pone en evidencia que los comportamientos de riesgo están vinculados a normas aprendidas.

Por ello, la lucha feminista puede beneficiar a hombres, mujeres, y personas que se identifican fuera del binarismo de género. Sería beneficioso encontrar un nuevo modelo de masculinidad que sirva como referente para los jóvenes, no solo adaptándonos a la nueva realidad sino transformándola.

El objetivo de este trabajo es analizar la percepción de la masculinidad en los jóvenes de España con el fin de explorar los diferentes pensamientos y creencias que existen. Se pretende identificar no solo qué tipo de masculinidad reproducen los jóvenes, sino también si existen señales de cambio o evolución. Para abordar este objetivo se analizan distintos aspectos sociales y conductuales con el propósito de comprender cómo se relacionan con la percepción del concepto de masculinidad que tienen los jóvenes.

En primer lugar, se realiza un análisis clúster con el fin de segmentar los individuos según su grado de acuerdo o desacuerdo hacia la masculinidad hegemónica. Tras este análisis, se identifican tres perfiles, masculinidad tóxica, que perpetúan pensamientos tradicionales, masculinidad neutra, que muestra un perfil intermedio y masculinidad sana, perfil que rechaza los estereotipos masculinos tradicionales.

Para analizar qué factores predicen la pertenencia a cada uno de estos perfiles, se realizan modelos explicativos. Se aplican modelos de regresión logística multinomial, con diversos métodos de selección de variables, que permite identificar relaciones lineales de variables sociodemográficas, actitudinales y de comportamiento. Además, se crea un árbol de decisión para capturar relaciones no lineales o jerárquicas, que ofrece una representación visual.

Ambos enfoques se complementan para ofrecer una visión completa de los factores asociados a los distintos modelos de masculinidad en la juventud española.

## 3. Metodología

### 3.1. Algoritmo MissForest

El algoritmo missForest es una técnica no paramétrica propuesta por Stekhoven y Bühlmann (2012) para la imputación de valores perdidos. Este procedimiento hace uso del *Random Forest* (Breiman, 2001), lo que permite trabajar con todo tipo de variables, tanto cuantitativas como cualitativas.

El algoritmo missForest tiene como objetivo predecir los valores faltantes de manera iterativa hasta alcanzar la convergencia, manteniendo de esta manera, todas las observaciones de los datos. Se aprovecha la robustez y precisión de la herramienta *Random Forest*, y aporta las estimaciones de tasas de error *out-of-bag* (OOB), lo que permite evaluar el procedimiento sin la necesidad de contar con un conjunto de datos para realizar dicha evaluación.

Sea un conjunto de datos  $X = (X_1, X_2, \dots, X_p)$ , una matriz de datos de dimensión  $n \times p$ , donde existen datos faltantes dispersos en las distintas variables. Teniendo en cuenta los valores perdidos de una variable arbitraria  $X_s$  en las posiciones  $i_{mis}^{(s)} \subseteq \{1, 2, \dots, n\}$ , se define:

1.  $y_{obs}^{(s)}$ : los valores observados de la variable  $X_s$ .
2.  $y_{mis}^{(s)}$ : los valores perdidos de la variable  $X_s$ .
3.  $x_{obs}^{(s)}$ : variables distintas de  $X_s$  con observaciones  $i_{obs}^{(s)} = \{1, 2, \dots, n\} \setminus i_{mis}^{(s)}$ .
4.  $x_{mis}^{(s)}$ : variables distintas de  $X_s$  con observaciones  $i_{mis}^{(s)}$ .

El procedimiento missForest consiste en los siguientes pasos:

1. Se realiza una primera imputación por estimaciones simples, como la media, generando  $X^{imp}$ .
2. Se ordenan las variables con valores perdidos, comenzando por aquellas que tienen menos valores faltantes.
3. Para cada variable  $X_s$  se ajusta un modelo *Random Forest* utilizando  $y_{obs}^{(s)}$  como variable respuesta y  $x_{obs}^{(s)}$  como variables predictoras.
4. Se predicen los valores perdidos aplicando el *Random Forest* a  $x_{mis}^{(s)}$  y se sustituyen los nuevos valores a los anteriores, generando  $X_{new}^{imp}$ .
5. Se repiten los pasos 3 y 4 hasta que se verifica el criterio de parada. Dicho criterio define que el algoritmo finaliza cuando la diferencia entre dos matrices de imputación consecutivas aumenta simultáneamente para variables continuas y categóricas. Esta condición se evalúa mediante dos medidas:

- Para variables continuas:

$$\Delta_N = \frac{\sum_{j \in N} (X_{new}^{imp} - X_{old}^{imp})^2}{\sum_{j \in N} (X_{new}^{imp})^2}$$

- Para las variables categóricas:

$$\Delta_F = \frac{\sum_{j \in F} \sum_{i=1}^n I_{X_{new}^{imp} \neq X_{old}^{imp}}}{\#NA}$$

Donde #NA es el número de valores perdidos en las variables categóricas.

Tras la imputación de los datos faltantes, se evalúa el rendimiento de la imputación. Se realiza mediante dos métricas:

- Error cuadrático medio normalizado (NRMSE), utilizado para variables continuas y se define como:

$$NRMSE_j = \sqrt{\frac{\sum_{i \in i_{mis}^{(j)}} (x_{ij}^{imp} - x_{ij}^{true})^2}{\sum_{i \in i_{mis}^{(j)}} (x_{ij}^{true})^2}}$$

Donde  $x_i^{imp}$  es el valor imputado, y  $x_i^{true}$  es el valor verdadero.

- Proporción de clasificación errónea (PCF), aplicable a variables categóricas que se calcula como:

$$PCF_j = \frac{1}{\#NA_j} \sum_{i \in i_{mis}^{(j)}} \mathbb{I}(x_i^{imp} \neq x_i^{true})$$

Donde  $\mathbb{I}$  es la función indicadora que toma valor 1 si la condición se cumple y 0 si no.

Para facilitar la comparación de errores, se procede a dividir cada error de las variables continuas por la desviación típica de su respectiva variable. Esta estandarización permite comparar e identificar aquellas variables cuya imputación ha sido imprecisa.

### 3.2. Matriz correlación de Pearson

Se realiza el cálculo de la matriz de correlaciones de Pearson con objetivo exploratorio. Este análisis permite identificar el grado de asociación lineal entre pares de variables cuantitativas.

La correlación de Pearson mide la fuerza y dirección de una relación lineal entre dos variables y su valor oscila entre -1 y 1. Valores cercanos a 1, indican una alta correlación lineal positiva, mientras que un valor cercano a -1 indica alta correlación negativa. Los valores cercanos a 0 indican ausencia de relación.

Se visualiza la matriz en un gráfico de correlación que permite detectar patrones entre las variables.

### 3.3. Análisis Clúster

El objetivo principal del análisis clúster es identificar distintos grupos, donde los datos dentro de cada grupo poseen características similares, y a su vez, presentan diferencias con el resto de los grupos.

Se hace uso del algoritmo k-means, se trata de una técnica que busca dividir las observaciones en k clústeres  $C = \{C_1, C_2, \dots, C_k\}$ , minimizando la varianza intra-clúster. Este proceso sigue los siguientes pasos:

1. Se seleccionan aleatoriamente k centroides iniciales  $\mu_1, \dots, \mu_k$ .
2. Cada observación se asigna al clúster cuyo centroide es más cercano a esta usando la distancia euclídea.

$$d(x_i, \mu_j) = \sqrt{\sum_{l=1}^n (x_{il} - \mu_{jl})^2}$$

3. Se calculan nuevos centroides como la media de las observaciones asignadas a cada clúster.
4. Se repiten los pasos 2 y 3 hasta que las asignaciones de clúster no cambian o se alcanza un número máximo de iteraciones.

Previo a este proceso, se debe seleccionar el valor  $k$  óptimo. Dicha elección se realiza mediante el método del codo. Este método se basa en evaluar la variación interna de los grupos, a medida que aumenta el valor de  $k$ . Con este fin, se calcula la suma total de errores cuadráticos para diferentes valores de  $k$  y se grafican.

$$SSE(k) = \sum_{i=1}^k \sum_{x \in C_i} (x - \mu_i)^2$$

Observando el gráfico, el “codo” de la curva, es decir, el punto donde la reducción del error comienza a ser más pequeña, sugiere el número óptimo de clústeres.

### 3.3.1. Reducción dimensionalidad: Análisis Componentes Principales

El objetivo del Análisis de Componentes Principales (PCA) es reducir la dimensionalidad de un conjunto de datos manteniendo la mayor parte de la varianza de los datos. Se usa dicho método para representar gráficamente los clústeres creados y facilitar su interpretación.

Para lograrlo, se obtienen nuevas variables, llamadas componentes principales, que son combinaciones lineales de las variables originales. Dichos componentes se ordenan según la varianza explicada, siendo el primer componente el que más información proporciona. A continuación, se seleccionan los dos primeros componentes para representar los clústeres en un gráfico bidimensional.

El gráfico resultante permite observar la separación y solapamiento de los grupos generados por el clustering. Este análisis permite caracterizar cada grupo.

### 3.4. Regresión logística multinomial

La regresión logística multinomial es una ampliación del modelo logístico binario. Su objetivo es estimar la probabilidad de que una observación pertenezca a una de las  $K$  categorías cualitativas de la variable dependiente  $Y$ , en función de un conjunto de variables independientes  $X_1, X_2, \dots, X_m$ , siendo  $K > 2$ .

A diferencia del modelo logístico binario donde se estima la probabilidad de que un evento ocurra o no, en el multinomial se elige una categoría de referencia, y el modelo estima  $K - 1$  funciones logísticas del tipo:

$$\log \left( \frac{P(Y = j | \mathbf{X})}{P(Y = K | \mathbf{X})} \right) = \beta_{0j} + \beta_{1j}X_1 + \dots + \beta_{mj}X_m$$

donde:

- $j = 1, \dots, K - 1$ .
- $\mathbf{X}$  es el vector de predictores (variables independientes).
- $\beta_{ij}$  son los coeficientes de cada categoría  $j$  respecto a la de referencia.

La función de predicción para cada categoría  $j$  es:

$$P(Y = j | \mathbf{X}) = \frac{e^{\beta_{0j} + \sum \beta_{ij} X_i}}{1 + \sum_{l=1}^{K-1} e^{\beta_{0l} + \sum \beta_{il} X_i}}$$
$$P(Y = 1 | \mathbf{X}) = \frac{1}{1 + \sum_{l=1}^{K-1} e^{\beta_{0l} + \sum \beta_{il} X_i}}$$

Se observa que esta formulación del modelo proporciona  $K - 1$  conjuntos de parámetros  $\beta$ , para todas las categorías, exceptuando la de referencia.

### 3.4.1. Métodos de selección de variables

Se hace uso de 3 métodos de selección de variables:

- Backward o hacia atrás: parte del modelo que contiene todos los posibles efectos. Se van eliminando una a una las variables que menos influyen en el modelo hasta que este empeora con la eliminación de cualquiera de las restantes. Una vez que un efecto se elimina, no vuelve a introducirse en el modelo.
- Forward o hacia delante: parte del modelo de cero, se van introduciendo uno a uno los efectos que mayor mejora producen hasta que no hay ningún efecto fuera del modelo que aporte información. Una vez un efecto entra en el modelo, no puede salir.
- Stepwise o paso a paso: este método es similar al método Forward, sin embargo, tiene la ventaja de que se pueden eliminar los efectos que han entrado en el modelo, si la entrada de otro hace que no sea significativa la aportación que este proporcionaba. La eliminación de los efectos se hace de acuerdo con el método Backward. Esto da lugar, a que en cada paso se evalúan los efectos a eliminar y a introducir y se realiza aquella acción que mayor mejora produzca.

Para evitar el sobreajuste del modelo, se recurre a unas medidas que evalúan la mejora producida en la verosimilitud, a la vez que tienen en cuenta la complejidad del modelo. Estas medidas son:

- AIC (Akaike information criterion):  $-2 \ln(L) + 2\tau$
- BIC (Bayesian information criterion):  $-2 \ln(L) + \tau \ln(n)$

La única diferencia entre estas medidas es la penalización al número de parámetros ( $\tau$ ), siendo el criterio BIC el que más penaliza y, por tanto, dará lugar a modelos menos complejos.

### 3.4.2. Interpretación de los parámetros

Los parámetros estimados se interpretan a través de los *odds ratio* (OR), es decir, los cocientes de las probabilidades. Los OR representan cómo cambian las probabilidades de estar en una categoría frente a la de referencia. Un OR mayor que 1 indica un aumento en la probabilidad de una categoría frente a la de referencia y un OR menor que 1, lo contrario.

## 3.5. Árbol de decisión

Los árboles de decisión son una herramienta útil para la predicción de variables cuantitativas y cualitativas realizando una segmentación de los datos de forma jerárquica y secuencial, obteniéndose de esta manera segmentos denominados nodos que representan submuestras del conjunto de datos.

Se implementa un árbol de clasificación ya que la variable objetivo es de tipo cualitativa. Este modelo permite representar de manera gráfica y jerárquica la clasificación de las observaciones en las diferentes categorías de la variable dependiente. La estructura del árbol consta de:

- Nodo raíz: contiene el 100% de la muestra.
- Nodos intermedios: representan reglas de partición sobre las variables explicativas.
- Nodos terminales u hojas: asignan una categoría de la variable dependiente.

Para la construcción del árbol de clasificación se construye un primer árbol inicial denominado maximal. Durante la construcción de dicho árbol, en cada nodo, el algoritmo evalúa todas las variables posibles y se selecciona aquella que mejor divide los datos según el índice de Gini. Esta medida permite establecer cómo de homogéneas son las observaciones de cada nodo con respecto a la variable objetivo.

$$IG(\text{nodo}_j) = 1 - \sum_{k=1}^K (p_{jk})^2$$

donde  $p_{jk}$  es la proporción de observaciones de la categoría  $k$  en el nodo  $j$ .

Para evitar el sobreajuste, se aplica una poda del árbol maximal, es decir, se eliminan divisiones que aportan poca información. Para realizar dicha poda se define el riesgo de los subárboles  $R_\alpha(T)$  y se busca aquel que minimiza dicha cantidad. Este criterio evalúa el error cometido por el subárbol, además de penalizar el tamaño del mismo ( $\alpha$ ).

$$R_\alpha(T) = \text{error}(T) + |T| \alpha \text{error}(T_1)$$

Se genera una secuencia de los subárboles asociados a los valores críticos ( $\alpha$ ), se evalúa mediante validación cruzada que permite estimar el error de clasificación. Después, se calcula el área bajo la curva ROC (AUC) como criterio adicional para la selección del mejor árbol. Esta medida evalúa la capacidad del modelo para distinguir entre categorías, su valor oscila de 0 a 1, indicando una mayor calidad cuánto más cercano a 1 sea.

El árbol obtenido final permite identificar combinaciones de distintas variables que llevan a una categoría determinada. Además, el modelo permite realizar un ranking de importancia de las distintas variables usadas según la mejora del índice de Gini que estas producen.

### 3.6. Monte Carlo Cross-Validation (MCCV)

Con el fin de comparar los diferentes modelos de regresión logística, se aplica la estrategia de validación conocida como Monte Carlo Cross-Validation (MCCV), también denominada validación cruzada repetida con particiones aleatorias.

MCCV consiste en realizar múltiples divisiones aleatorias del conjunto de datos total en dos subconjuntos. El primero es el conjunto de entrenamiento, con el 80% de la muestra y el conjunto prueba con el 20% restante, permitiendo de esta manera, generar múltiples combinaciones de entrenamiento y prueba.

Tras cada división, se entrenan todos los modelos sobre el conjunto de entrenamiento y se evalúan sobre el conjunto de prueba utilizando el Área Bajo la Curva ROC (AUC), que permite medir la capacidad del modelo para discriminar correctamente entre las clases. El procedimiento se repite durante 20 iteraciones. Para cada modelo, se calculan los

promedios de AUC y se visualizan los resultados en un gráfico boxplot, lo que permite comparar la estabilidad y el rendimiento medio de los distintos modelos visualmente.

El AUC mide la capacidad del modelo para distinguir entre categorías, su valor varía entre 0 y 1, donde un AUC igual a 0.5 indica un rendimiento equivalente al azar, y superior a 0.7 se considera buen modelo discriminante.

Gracias al uso de esta estrategia, se consigue mayor robustez frente a la aleatoriedad de una única partición de validación, sin embargo, tiene la desventaja de tener fuerte carga computacional.

Tras la aplicación de esta estrategia, se selecciona el mejor modelo de regresión logística multinomial, aquel que muestra mayor capacidad discriminante (mayor AUC promedio) y menor variabilidad entre iteraciones, además, permite comparar dicho modelo con el árbol de decisión, proporcionando una visión integral sobre qué enfoque ofrece mejores resultados.

## 4. Depuración de datos

### 4.1. Lectura de datos

La base de datos utilizada proviene del Centro Reina Sofía de Adolescencia y Juventud de la encuesta *La caja de la masculinidad: construcción, actitudes e impacto en la juventud española*.

El trabajo de campo se realizó en mayo del 2022 mediante una encuesta online con cuestionario cerrado. Cuenta con una muestra de 1687 individuos de entre 15 a 29 años, residentes en España. El estudio se basa en un muestreo estratificado con afijación proporcional, en el que se establecieron cuotas aleatorias en función del género (mujeres, hombres y otras identidades) y de tres grupos de edad (15-19, 20-24 y 25-29 años). Los datos han sido ponderados según el nivel educativo, agrupado en tres categorías: estudios obligatorios, secundarios postobligatorios y estudios superiores. Bajo el supuesto de un muestreo aleatorio simple (MAS), con máxima heterogeneidad y un nivel de confianza del 95%, el error muestral estimado es de  $\pm 2,36\%$ .

La encuesta se divide en 5 bloques temáticos y de las 238 variables de la encuesta se seleccionan 53 para el estudio, las cuales se han renombrado para facilitar su interpretación y análisis.

A continuación se muestran las variables que serán usadas en el estudio divididas en los 5 bloques de la encuesta.

#### I. Datos Personales

Tabla 1: Clasificación variables 1

Variable	Etiqueta	Tipo
EDAD	Edad	Numérica
GENERO	Género	Catagórica
CLASE_SOCIAL	Clase Social auto percibida	Catagórica
RELIGIOSIDAD	¿Podrías decir cuál es tu grado de religiosidad independiente de la religión?	Escala (0-10)
IDEOLOGIA	En una escala de 0 a 10, siendo el 0 la “extrema izquierda” y el 10 la “extrema derecha”, ¿dónde te situarías?	Escala (0-10)
ORIENTACION	Dirías que tu orientación sexual es:	Catagórica

#### II. Bienestar y Situación Personal

Tabla 2: Clasificación variables 2

Variable	Etiqueta	Tipo
FELICIDAD	De forma general, considerando la satisfacción con todos los aspectos de tu vida, ¿cómo valorarías tu grado de felicidad en el momento actual?	Escala (0-10)
TU_MISMIX	En tu vida diaria, ¿con qué frecuencia piensas que eres realmente tú mismo/a?	Escala (0-10)
SALUD_MENTAL	En los últimos doce meses, ¿has tenido o has creído tener alguna vez un problema psicológico, psiquiátrico o de salud mental?	Catagórica

HABLAR_AMIGXS	¿Tienes amigos o amigas con los que te sientas cómodo/a hablando sobre temas personales y emocionales en tu vida?	Categoría
---------------	---	-----------

### III. Personalidad y Apariencia

Las siguientes variables se refieren a la presión que sienten los individuos en los siguientes aspectos:

Tabla 3: Clasificación variables 3

Variable	Etiqueta	Tipo
P_EXITO	Tener éxito en el trabajo o estudio	Escala (0-10)
P_GRUPO	Actuar o pensar como mi grupo de amigos/as	Escala (0-10)
P_REDES	Tener éxito en las redes sociales (obtener comentarios positivos o “me gusta”, conseguir seguidores...)	Escala (0-10)
P_EMOCIONES	Ocultar la tristeza o la ansiedad	Escala (0-10)
P_APARIENCIA	Ser físicamente atractivo o atractiva	Escala (0-10)
P_SEXUAL	Ser bueno/a practicando sexo	Escala (0-10)
P_HOGAR	Realizar las tareas del hogar	Escala (0-10)
P_APOYO	Cuidar y apoyar emocionalmente a otras personas	Escala (0-10)
P_VIOLENCIA	Usar la violencia para defender mi reputación o ideas	Escala (0-10)
P_REL_HETERO	Relacionarme exclusivamente con personas heterosexuales	Escala (0-10)
P_FORTALEZA	Mostrar fortaleza, aunque esté nervioso/a o asustado/a	Escala (0-10)
P_ALC_DROGAS	Consumir alcohol u otras drogas en contextos de ocio	Escala (0-10)
P_PAREJA	Tener pareja estable	Escala (0-10)
P_LIGAR	Llevar la iniciativa a la hora de ligar	Escala (0-10)

A continuación, se muestran variables referentes a acciones de riesgo. ¿Ha realizado el individuo estas acciones en los últimos 6 meses?:

Tabla 4: Clasificación variables 4

Variable	Etiqueta	Tipo
CON_ALCOH_1	He conducido después de haber bebido alcohol	Dicotómica
CON_DROG_2	He conducido habiendo consumido cannabis, cocaína, speed o pastillas	Dicotómica
CON_VELOC_3	He conducido a mucha más velocidad de la permitida	Dicotómica
SEX_RIESGO_4	He tenido conductas sexuales de riesgo (sin preservativo, chemsex...)	Dicotómica
PELEAS_5	He participado en peleas	Dicotómica
COCHE_PELIGRO_6	He ido en un coche en el que la persona que conducía iba bebida/drogada	Dicotómica
INDICE_RIESGO	Acciones de riesgo que ha realizado el individuo	Escala (0-6)

Se ha creado una nueva variable llamada INDICE\_RIESGO, que indica el número de acciones de riesgo que ha realizado el individuo con las variables CON\_ALCOH\_1, CON\_DROG\_2, CON\_VELOC\_3, SEX\_RIESGO\_4, PELEAS\_5, COCHE\_PELIGRO\_6. De manera que la nueva variable puede tomar valores de entre 0 a 6.

Por último, se muestran variables referentes a la frecuencia con la que el individuo ha realizado las siguientes acciones:

Tabla 5: Clasificación variables 5

Variable	Etiqueta	Tipo
BURLA	Te burlaste de alguien o llamaste a alguien de alguna manera que no le gustó, por alguna razón	Escala (1-5)
AVERGONZAR	Insultaste a alguien, publicaste fotos para avergonzarlo/a, o le amenazaste por mensajes de texto, redes sociales u otra aplicación o página web	Escala (1-5)
VIOLENCIA	Ejerciste violencia física contra alguien de forma deliberada en mayor o menor medida	Escala (1-5)

#### IV. La Masculinidad

Tabla 6: Clasificación variables 6

Variable	Etiqueta	Tipo
H_VERDAD_ENTORNO	Desde la infancia, ¿te han dicho de forma habitual desde tu entorno cercano (familia, amigos/as, vecinos, etc.) que un “hombre de verdad” se comporta o se tiene que comportar de cierta manera?	Dicotómica
M_VERDAD_ENTORNO	Desde la infancia, ¿te han dicho de forma habitual desde tu entorno cercano (familia, amigos/as, vecinos, etc.) que una “mujer de verdad” se comporta o se tiene que comportar de cierta manera?	Dicotómica

A continuación, se muestran las variables referentes a la percepción de la masculinidad. Independientemente de la sociedad, ¿hasta qué punto está de acuerdo el individuo con estas afirmaciones?:

Tabla 7: Clasificación variables 7

Variable	Etiqueta	Tipo
MASC_DEBIL_DEFENSA	Un hombre que no se defiende cuando otros abusan de él, es débil	Escala (1-4)
MASC_RESUELVE_SOLO	Los hombres deberían resolver sus problemas personales por sí mismos sin pedir ayuda a los demás	Escala (1-4)
MASC_LLORAR	Un hombre no debería llorar en público	Escala (1-4)
MASC_VIOL_RESP	Es lógico que los hombres usen la violencia para obtener respeto si es necesario	Escala (1-4)

MASC_TAR_CASA	No es bueno que se enseñe a un niño cómo cocinar, coser, limpiar la casa y cuidar a niños más pequeños	Escala (1-4)
MASC_PROVEEDOR_HOGAR	Los hombres deberían ser realmente los que traigan dinero al hogar y sostengan a sus familias, no las mujeres	Escala (1-4)
MASC_DECISION_REL	Un hombre siempre debería tener la última palabra sobre decisiones en su relación o matrimonio	Escala (1-4)
MASC_JUEG_GENERO	Un niño debe evitar jugar desde pequeño a “cosas de niñas”	Escala (1-4)
MASC_AFEMINADO	Un hombre de verdad nunca puede ser afeminado	Escala (1-4)
MASC_PROMISCUO	Un hombre debería tener la mayor cantidad de parejas sexuales que pueda	Escala (1-4)
MASC_INFIDELIDAD	Si un hombre no tiene suficiente sexo con su pareja es comprensible que cometa una infidelidad	Escala (1-4)
MASC_PROST_NORMAL	Es normal que un hombre recurra a la prostitución para tener sexo	Escala (1-4)
MASC_PIROPOS_NORM	Es normal que un hombre piropoee a una mujer que no conoce en un espacio público si le parece atractiva	Escala (1-4)
MASC_AMIGOS_HOMO	No es aconsejable que los hombres heterosexuales tengan amigos homosexuales	Escala (1-4)
MASC_VERD_HETERO	Los hombres de verdad son heterosexuales	Escala (1-4)
MASC_MUJ_CUIDADO	Las mujeres no buscan hombres que se preocupan demasiado por su aspecto y forma de vestir	Escala (1-4)
MASC_NO_CUIDADO	Un hombre que pase mucho tiempo ocupándose de su apariencia no es muy masculino	Escala (1-4)

## V. Relaciones, Roles y Creencias

Tabla 8: Clasificación variables 8

Variable	Etiqueta	Tipo
FEMINISTA	¿Te consideras feminista?	Dicotómica

### 4.2. Validación NS/NC

En la base de datos no existen valores perdidos, sin embargo, se analizan las respuestas no sabe o no contesta (NS/NC) ya que, si un individuo ha respondido muchas preguntas con esta opción, no será apto para continuar con el estudio.

Recodificamos la categoría NS/NC (codificada como 99) como valor NA para analizarla.

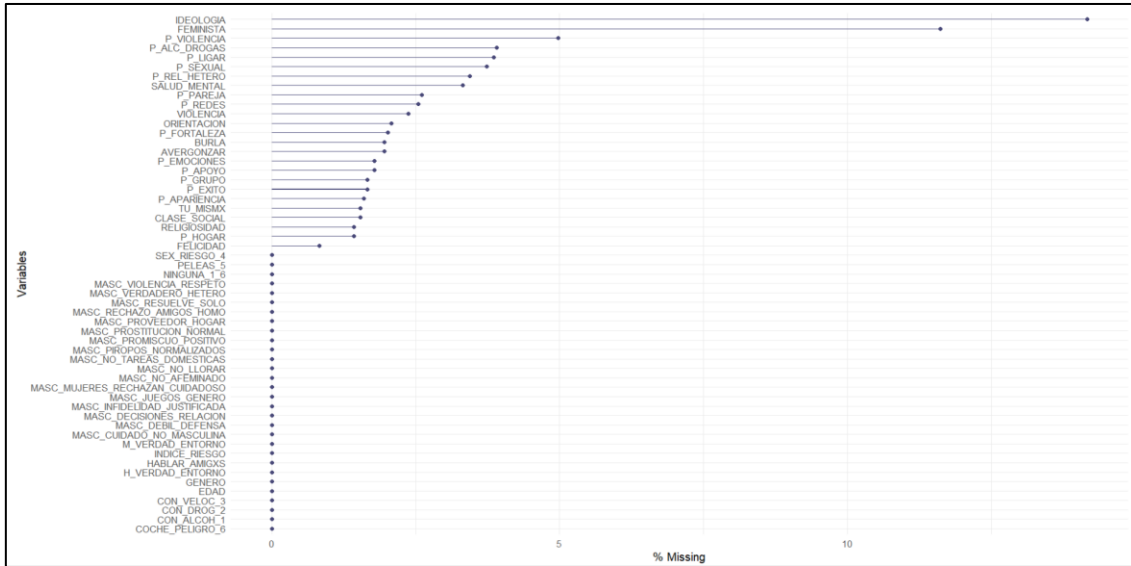


Figura 1: Valores NA por variables

Las variables con mayor porcentaje de NA son IDEOLOGÍA (con un 14% de valores NA) y FEMINISMO (con un 11% de valores NA). El resto de las variables presentan un porcentaje igual o inferior al 5% de valores NA.



Figura 2: Valores NA por observaciones

Se observa (Figura 2) que no hay un número elevado de individuos que hayan respondido la opción NS/NC en muchas preguntas, por lo que son válidas todas las observaciones.

En la variable FEMINISMO mantendremos la categoría NS/NC ya que, al tener un número elevado de observaciones con esta opción, resulta de interés.

### 4.3. Tratamiento valores NA

Con el fin de mantener todas las observaciones del conjunto de datos, se realiza una imputación de datos faltantes mediante el algoritmo missForest. Este método se basa en random forest, por lo que puede trabajar con variables tanto cualitativas como cuantitativas. Se realiza este proceso con un número de árboles igual a 150, para obtener buenos resultados sin una carga computacional excesiva.

Para evaluar la calidad de la imputación realizada, se calculan los errores de imputación específicos por variable, es decir, la dificultad que tuvo el modelo en predecir los valores de cada variable. Este error representa el error cuadrático medio (MSE) en variables numéricas, o la proporción de clasificación errónea (PFC) en variables categóricas. Dichos errores se proporcionan en la siguiente tabla incluyendo solo aquellas variables con errores distintos de 0:

Tabla 9: Errores imputación

Variable	Error	Variable	Error
CLASE_SOCIAL	0.5454	RELIGIOSIDAD	8.088
IDEOLOGIA	4.2112	ORIENTACION	0.2015
FELICIDAD	3.5521	TU_MISMIX	4.0524
SALUD_MENTAL	0.5597	P_EXITO	5.5801
P_GRUPO	5.4214	P_REDES	6.1172
P_EMOCIONES	6.206	P_APARIENCIA	6.3728
P_SEXUAL	7.205	P_HOGAR	6.9616
P_APOYO	7.025	P_VIOLENCIA	6.5576
P_REL_HETERO	5.6494	P_FORTALEZA	5.2108
P_ALC_DROGAS	6.1723	P_PAREJA	7.6043
P_LIGAR	6.9305	BURLA	0.4268
AVERGONZAR	0.2194	VIOLENCIA	0.26896

Dado que los errores absolutos no son comparables entre variables con diferentes escalas, se procede a calcular los errores relativos dividiéndolos por sus desviaciones típicas. De este modo, es posible identificar las variables cuya imputación es problemática.

Tabla 10: Errores de imputación normalizados

Variable	Error	Variable	Error
CLASE_SOCIAL	0.5454	RELIGIOSIDAD	2.58
IDEOLOGIA	1.8597	ORIENTACION	0.2015
FELICIDAD	1.5847	TU_MISMIX	1.7367
SALUD_MENTAL	0.5597	P_EXITO	1.892
P_GRUPO	1.809	P_REDES	1.969
P_EMOCIONES	1.989	P_APARIENCIA	2.03
P_SEXUAL	2.1798	P_HOGAR	2.2668
P_APOYO	2.2484	P_VIOLENCIA	2.1524
P_REL_HETERO	1.8557	P_FORTALEZA	1.77
P_ALC_DROGAS	2.066	P_PAREJA	2.298
P_LIGAR	2.1259	BURLA	0.4268
AVERGONZAR	0.2194	VIOLENCIA	0.26896

Tras este análisis, se identifican algunos errores relativos ligeramente altos. Por ejemplo, RELIGIOSIDAD, P\_PAREJA, P\_SEXUAL, P\_APOYO, P\_HOGAR, P\_VIOLENCIA, P\_LIGAR, P\_ALC\_DROGAS Y P\_APARIENCIA. Estos errores pueden deberse a una alta dispersión de respuestas o baja correlación con otras variables del estudio.

A pesar de ello, se decide mantener todas las variables en el conjunto de datos, sin embargo, será necesario tratar con cautela los resultados derivados de las variables que muestran mayor dificultad de imputación.

## 5. Análisis Descriptivo

### 5.1. Perfil sociodemográfico

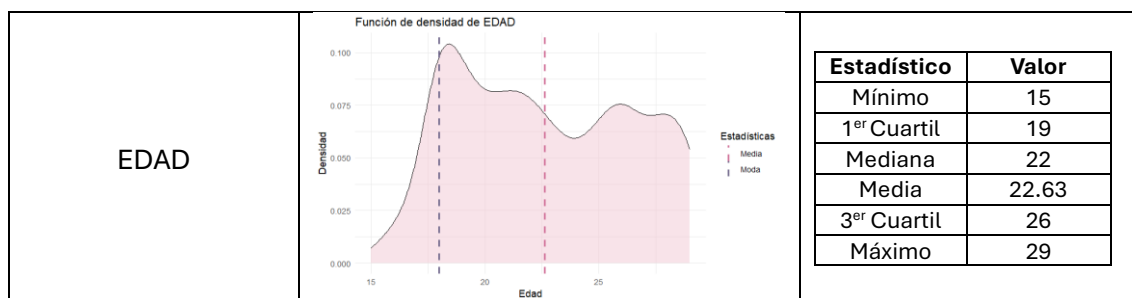


Figura 3: Distribución por edad

Los individuos encuestados son jóvenes que tienen edades desde los 15 hasta los 29 años, siendo la media global igual a 22.63.

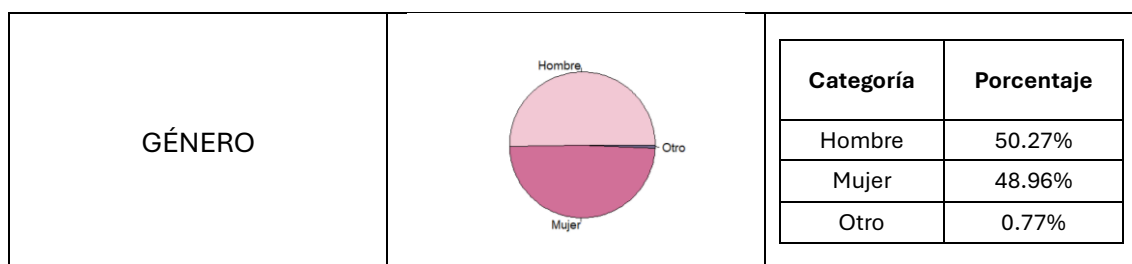


Figura 4: Distribución por género

En la base de datos existe aproximadamente la misma proporción de hombres que de mujeres con un 50.27% y 48.96% respectivamente. El 0.77% restante pertenece a individuos que se identifican con otro género. Al ser un porcentaje muy pequeño se valora eliminar dicha categoría, sin embargo, se mantiene por ser de interés para el estudio a pesar de tener escasa proporción.

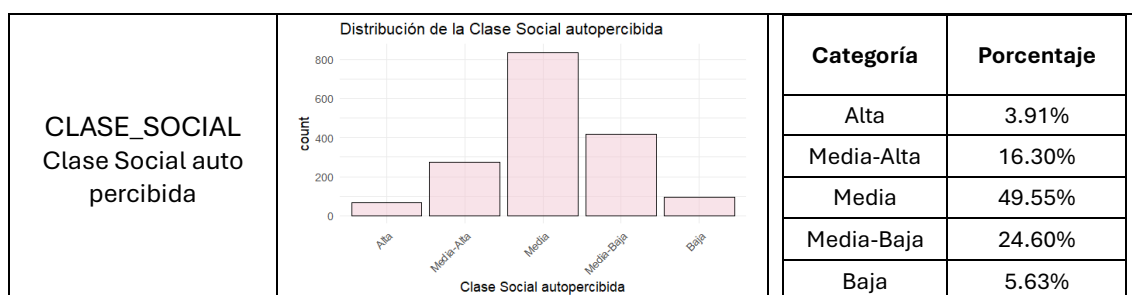


Figura 5: Distribución por clase social

La categoría mayoritaria de la variable CLASE\_SOCIAL es la clase media con casi el 50% de los individuos, seguida de la clase social media-baja con un 25% aproximadamente. Por lo tanto, los individuos se auto perciben mayoritariamente como de clase social media o media-baja.

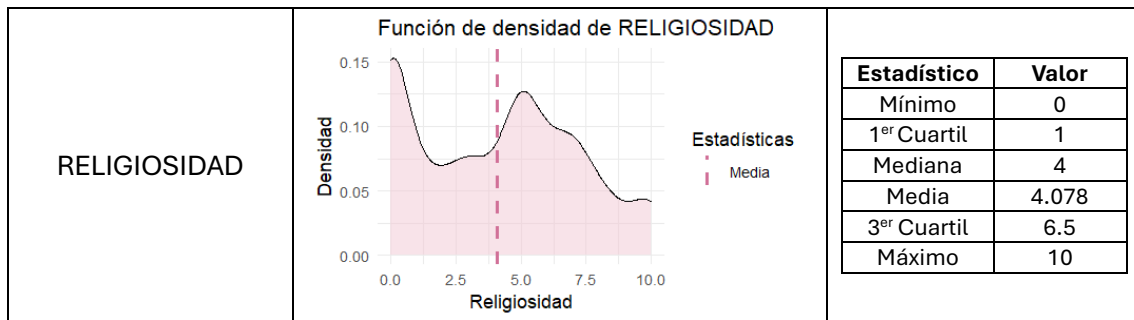


Figura 6: Distribución por religiosidad

La variable RELIGIOSIDAD tiene valores desde 0 a 10, siendo 0 “nada religiosa/o” y 10 “totalmente religiosa/o” independientemente de la religión. Se observa que hay mayor proporción de individuos que respondieron con valores cercanos a 0 siendo la media igual a 4 aproximadamente. Se tiene en cuenta que esta variable tuvo un error ligeramente alto en el proceso de imputación.



Figura 7: Distribución por ideología

La variable IDEOLOGÍA tiene valores desde 0 a 10, siendo 0 “extrema izquierda” y 10 “extrema derecha”. La mayor parte de los individuos se posicionan en el centro siendo la media igual a 4.862.

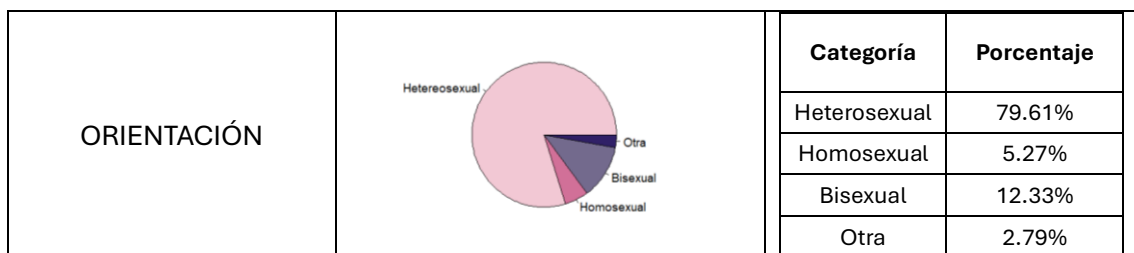


Figura 8: Distribución por orientación sexual

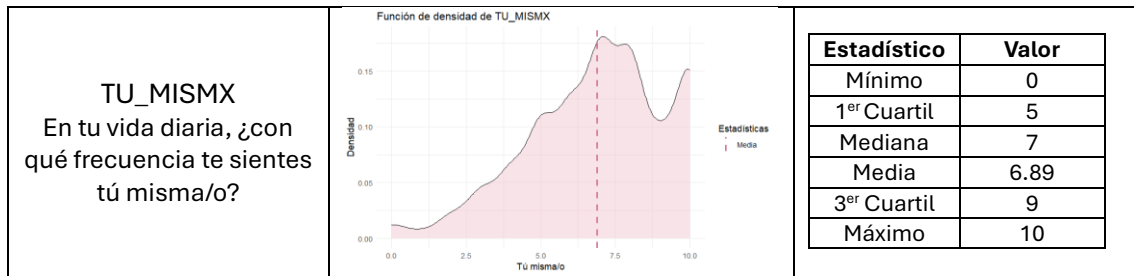
En la variable ORIENTACION la categoría mayoritaria es heterosexual con casi el 80% de los individuos. El resto de las categorías presentan porcentajes pequeños, destacando bisexual, la más representativa entre ellas con un 12.33%.

## 5.2. Bienestar y satisfacción personal

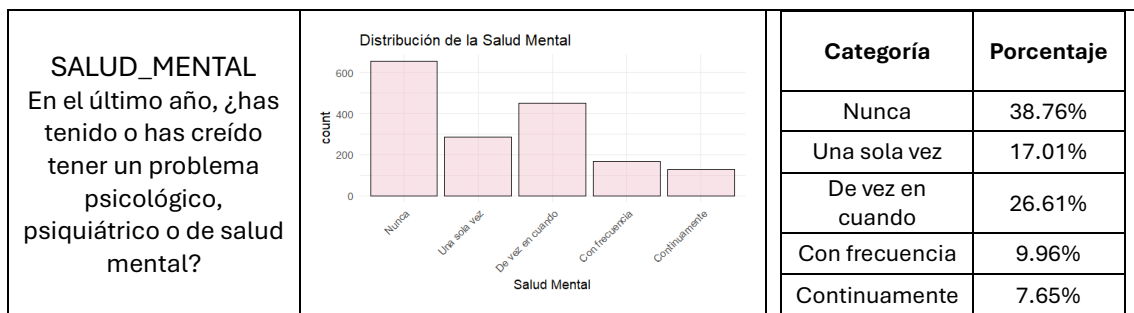


Figura 9: Distribución por estado de felicidad

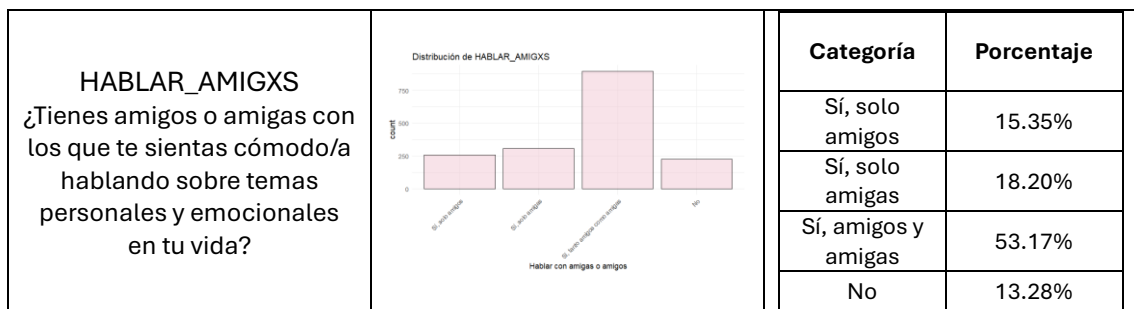
La variable FELICIDAD toma valores de entre 0 a 10, siendo 0 “totalmente infeliz” y 10 “totalmente feliz”. Se observa que la media se encuentra en el valor 6.55, es decir, mayoritariamente los individuos valoran que son felices.



La variable TU\_MISMX toma valores de entre 0 a 10, siendo 0 “nunca me siento yo misma/o” y 10 “todo el tiempo me siento yo misma/o”. Dicha variable toma de media el valor 6.89, es decir, los individuos sí se sienten ellos mismos con bastante frecuencia.



Casi el 40% de los individuos no han tenido ningún problema psicológico, psiquiátrico o de salud mental en el último año, sin embargo, un 7.65% de los individuos aseguran tener problemas psicológicos continuamente.



El 13% de los individuos afirma no tener amigos con los que se sientan cómodos para hablar sobre temas personales y el 53% afirma tener tanto amigos como amigas.

### 5.3. Presiones sociales y conductas de riesgo

A continuación, se observa la distribución de las variables que representan las presiones que sufren los jóvenes en distintos aspectos.



Figura 13: Distribución por presiones

También se muestra la media de respuestas de las diversas presiones evaluadas:

Tabla 11: Media de respuestas por presiones

Variable	Media
P_ALC_DROGAS	2.793
P_APARIENCIA	5.51
P_APOYO	4.934
P_EMOCIONES	5.486
P_EXITO	6.616
P_FORTALEZA	5.54
P_GRUPO	3.643
P_HOGAR	4.614
P_LIGAR	4.085
P_PAREJA	4.02
P_REDES	3.391
P_REL_HETERO	2.399
P_SEXUAL	4.517
P_VIOLENCIA	2.783

Se observa (Figura 13) que las variables P\_ALC\_DROGAS, P\_GRUPO, P\_LIGAR, P\_PAREJA, P\_REDES, P\_REL\_HETERO y P\_VIOLENCIA presentan asimetría positiva, lo que indica que los individuos sienten menor presión en estos ámbitos, es decir, consumir alcohol o drogas, actuar y pensar como su grupo de amigos, ligar o tener pareja, tener éxito en redes sociales, relacionarse exclusivamente con personas heterosexuales o usar la violencia para defender sus ideas.

Por otro lado, variables como P\_APARIENCIA, P\_EMOCIONES, P\_EXITO y P\_FORTALEZA presentan asimetría negativa. En consecuencia, los individuos sienten una mayor presión en ámbitos como la apariencia física, éxito en el trabajo, mostrar fortaleza en momentos de nerviosismo o miedo y ocultar la tristeza o la ansiedad.

Cabe añadir que, las variables P\_APOYO y P\_HOGAR muestran cierta simetría, con medias iguales a 4.936 y 4.615 respectivamente, por lo que los individuos sienten una presión moderada en estos aspectos.

Por último, se observa que el valor 0 presenta una frecuencia elevada en todas las variables, lo que indica que hay muchos individuos que no sienten nada de presión en los diversos ámbitos evaluados. La distribución que presentan estas variables, mayor frecuencia en el valor 0, es posible que sea el motivo por el cual se imputaron peor en el proceso de depuración.

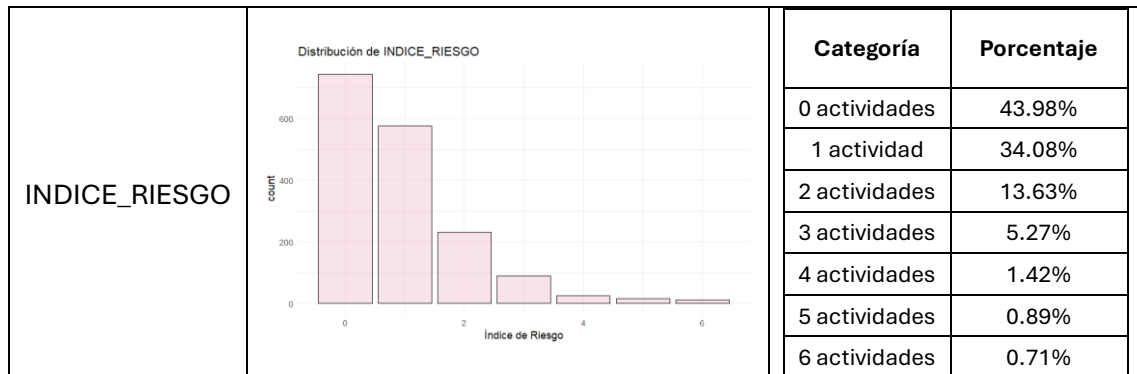


Figura 14: Distribución por índice de riesgo

Aproximadamente el 8% de los individuos han realizado 3 o más actividades de riesgo como conducir después de haber tomado alcohol o drogas, participar en peleas o tener conductas sexuales de riesgo. Casi el 44% de los individuos no han realizado ninguna de estas actividades en los 6 meses previos a la realización de la encuesta.

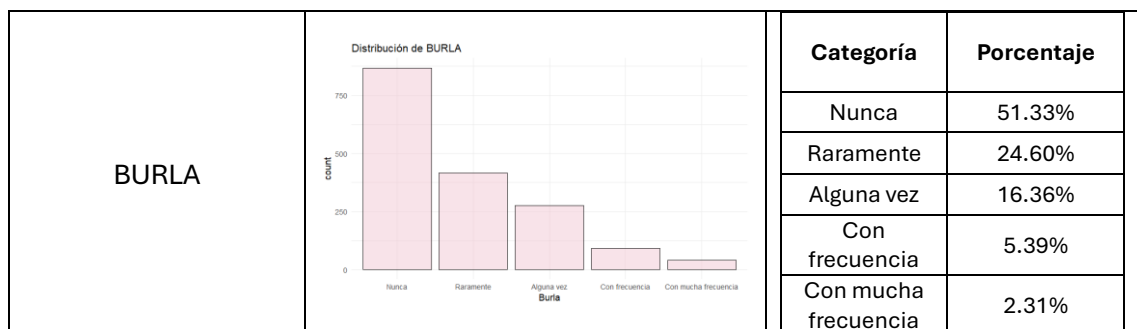


Figura 15: Distribución por frecuencia de burla

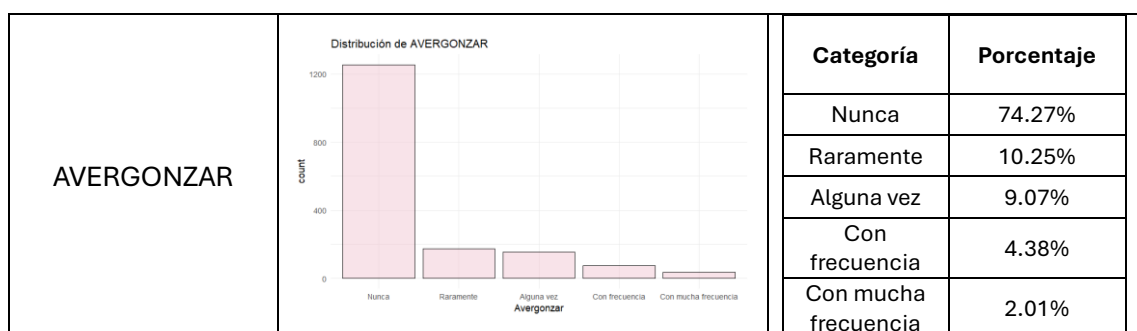


Figura 16: Distribución por frecuencia de avergonzar

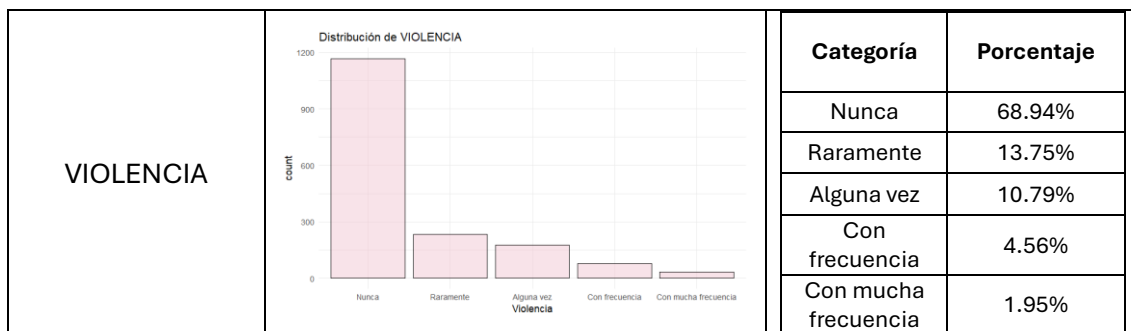


Figura 17: Distribución por frecuencia de violencia

Las variables BURLA, AVERGONZAR y VIOLENCIA, muestran una serie de comportamientos negativos y la frecuencia con la que los individuos los han realizado. Alrededor del 2% de los individuos realizan dichos actos con mucha frecuencia. El 51.39% asegura nunca haber hecho burla de alguien en los últimos 6 meses previos a la realización de la encuesta. El 74.39% afirma nunca haber avergonzado a alguien y por último el 69% asegura no haber ejercido violencia contra alguien.

#### 5.4. Entorno social

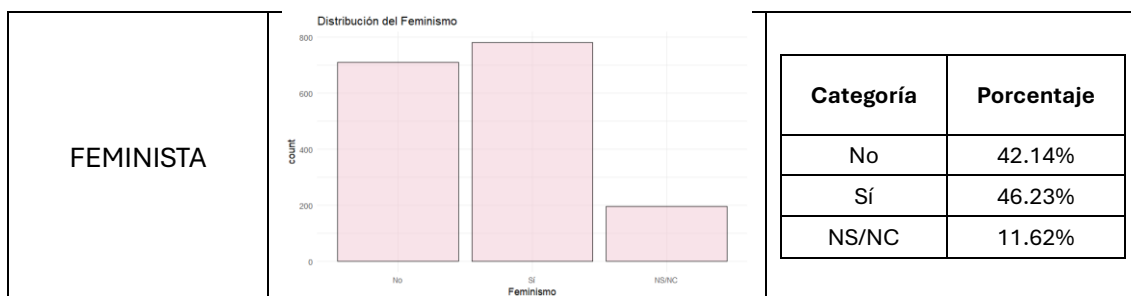


Figura 18: Distribución por feminismo

El 46.23% de los individuos afirman considerarse feminista frente a un 42.14% de individuos que afirmaron no serlo. Es importante destacar que casi el 12% de los individuos respondieron NS/NC a esta pregunta.

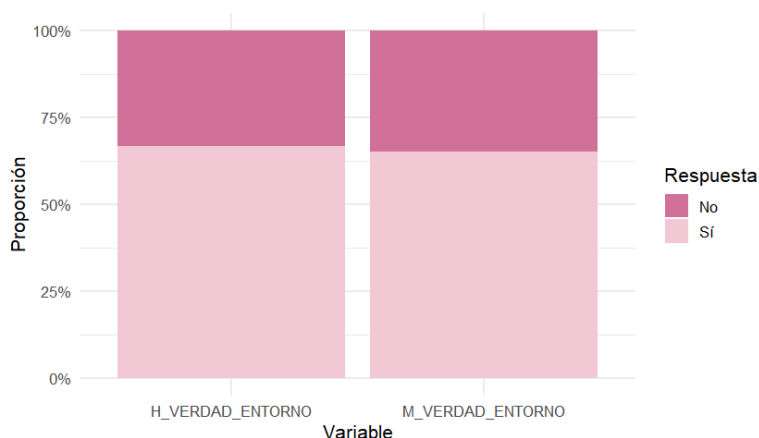


Figura 19: Distribución por transmisión roles de género

En cuanto a la transmisión de roles de género tradicionales, el 66.80% de los individuos han tenido un entorno familiar donde se les ha enseñado cómo se debe comportar un hombre de verdad. De igual forma, el 65.26% asegura haber tenido un entorno donde se le ha enseñado cómo se debe comportar una mujer de verdad. Prácticamente es el mismo

porcentaje siendo la pregunta con respecto a los hombres levemente superior. Esto refuerza la transmisión de roles de género tradicionales desde edades tempranas.

## 5.5. Percepciones de masculinidad

En los gráficos siguientes se muestra la distribución de las variables referentes a la percepción de la masculinidad.

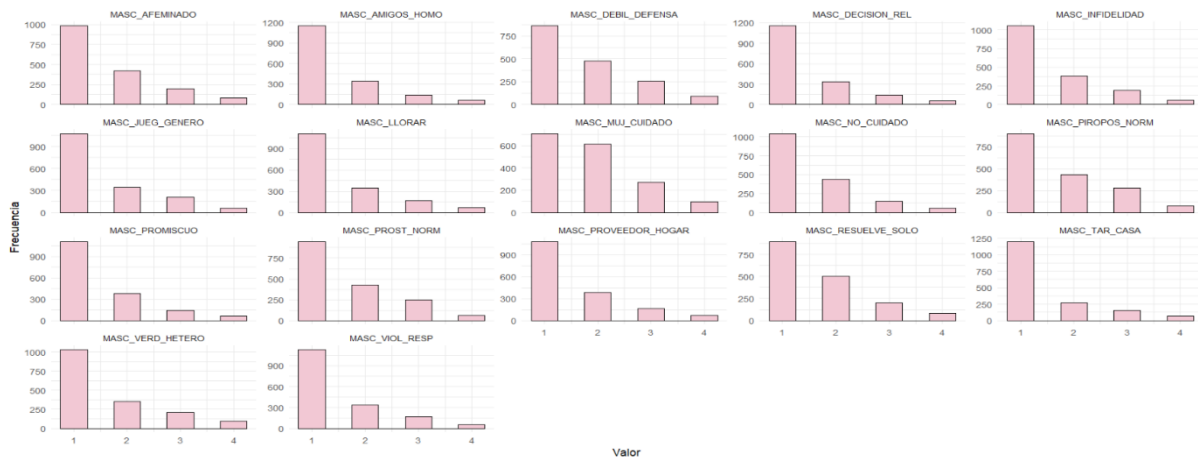


Figura 20: Distribución por percepción masculinidad

En todas las variables se observa (Figura 20) que la opción mayoritaria es el “1”, lo que indica que hay cierto rechazo hacia el estereotipo de la masculinidad tradicional entre los individuos encuestados. Este patrón indica una posible tendencia hacia una visión más igualitaria.

Sin embargo, hay una mayor dispersión en las variables MASC\_DEBIL\_DEFENSA, MASC\_MUJ\_CUIDADO, MASC\_PIROPOS\_NORM, MASC\_PROST\_NORM y MASC\_RESUELVE\_SOLO donde la opción “1” presenta una menor proporción que en el resto de las variables. La distribución más equilibrada en estas variables puede indicar una mayor simpatía o normalización por parte de los individuos hacia estas cuestiones.

Este contraste evidencia que, si bien existe un avance en la deconstrucción de algunos estereotipos de género, todavía persisten pensamientos tradicionales en ciertos ámbitos.

Como ya se ha señalado previamente, el objetivo de este estudio es analizar la percepción de la masculinidad por parte de los jóvenes. Es por ello por lo que se usarán las variables MASC, como variable objetivo en el análisis posterior.

Con un fin exploratorio, se calcula la matriz de correlaciones de Pearson entre las variables incluidas en este bloque, es decir, el grado de asociación lineal entre cada par de variables. El objetivo es encontrar relaciones entre ellas para facilitar el análisis posterior, además de detectar agrupaciones conceptuales.

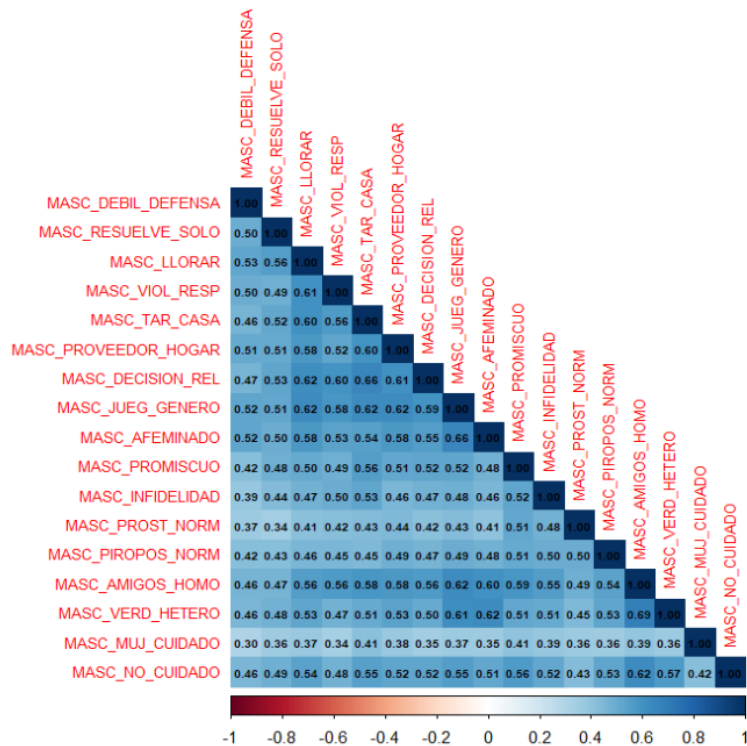


Figura 21: Matriz correlaciones Pearson de variables MASC

En primer lugar, se identifican correlaciones superiores a 0.6, lo que indica una fuerte asociación como en las variables MASC\_TAR\_CASA y MASC\_DECISION\_REL (0.66) o en las variables MASC\_AMIGOS\_HOMO y MASC\_VERD\_HETERO (0.69).

A su vez, se observan grupos temáticos entre variables con alta correlación, por ejemplo, las variables MASC\_JUEG\_GENERO, MASC\_AMIGOS\_HOMO, MASC\_VERD\_HETERO presentan correlaciones superiores a 0.6 entre sí y todas ellas relacionadas con la percepción del género, relaciones sociales y orientación sexual. Se observa un segundo grupo temático con las variables MASC\_TAR\_CASA, MASC\_PROVEEDOR\_HOGAR y MASC\_DECISION\_REL que representan liderazgo y dominio por parte del hombre en la relación de pareja y el hogar.

En conjunto, los resultados muestran un rechazo generalizado hacia la masculinidad tradicional en la mayoría de los ámbitos evaluados. Asimismo, el análisis de correlaciones sugiere que determinadas creencias y actitudes vinculadas a la masculinidad tienden a agruparse temáticamente. Estos hallazgos muestran una masculinidad en transformación, pero aún influida por normas tradicionales.

## 5.6. Otras relaciones

Es importante tener en cuenta la relación que existe entre las percepciones de la masculinidad y el género del individuo para tener una visión inicial de las diferencias existentes entre géneros.

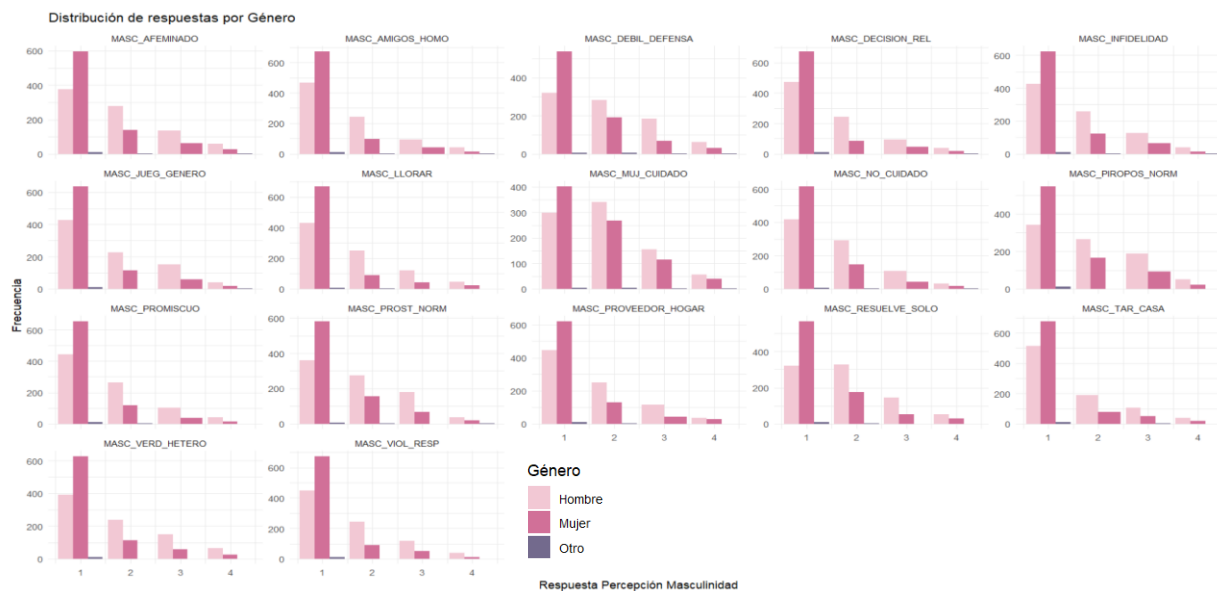


Figura 22: Distribución por percepción de masculinidad y género

Se aprecia (Figura 22) que el género mujer muestra un mayor desacuerdo con las afirmaciones sobre la masculinidad tradicional que el género hombre. Sin embargo, por lo general se observa un rechazo por parte de todos los individuos hacia la masculinidad tradicional lo que podría reflejar una evolución cultural, como se ha mencionado anteriormente. Este rechazo es especialmente evidente en las variables MASC\_TAR\_CASA (que no se enseñe a los niños cocinar, coser, limpiar la casa o cuidar a niños más pequeños) o MASC\_AMIGOS\_HOMO (relacionarse únicamente con personas heterosexuales), lo que podría indicar que en estos aspectos los individuos se alejan del pensamiento tradicional y muestran una evolución social.

Sin embargo, en las variables MASC\_MUJ\_CUIDADO (las mujeres no buscan hombres que se preocupan demasiado por su aspecto y forma de vestir), MASC\_PIROPOS\_NORM (la normalización de que un hombre piropoee a una mujer que no conoce en un espacio público) y MASC\_DEBIL\_DEFENSA (un hombre que no se defiende cuando otros abusan de él, es débil) hay una mayor proporción de aceptación lo que indica que en lo relativo a estas cuestiones, los individuos mantienen ese pensamiento tradicional.

En cuanto a otras identidades de género no binarias, se observa que muestran rechazo en la mayoría de las afirmaciones sobre la masculinidad tradicional. No obstante, hay una pequeña proporción que mantiene ese pensamiento tradicional en algunas cuestiones como MASC\_DEBIL\_DEFENSA, coincidiendo con las variables que tienen una mayor proporción de acuerdo. Por lo que refleja, que hay algunos aspectos donde persiste ese pensamiento tradicional sobre la masculinidad y los roles de género.

En resumen, aunque se observa que existe un avance social, ciertos discursos y actitudes asociadas a la masculinidad hegemónica aún se mantienen en la juventud, especialmente en los hombres.

## 6. Análisis Clúster

El objetivo principal del estudio es analizar la percepción de la masculinidad de los jóvenes. Para abordar dicho objetivo, se realiza un análisis clúster con el fin de agrupar a los individuos en distintos grupos, según sus respuestas.

Para este análisis se utiliza únicamente las variables referentes a la percepción de la masculinidad, ya que constituyen el eje central del objeto de estudio. De esta manera, se segmentará el conjunto de datos en distintos perfiles según su pensamiento hacia el concepto de masculinidad.

El método utilizado para realizar el análisis es k-means, una técnica de partición no jerárquica. Dicho método divide las observaciones en k grupos o clústeres. El objetivo es minimizar la varianza intra-clúster, garantizando que los individuos de un mismo grupo muestren perfiles similares, mientras que los distintos grupos muestran características diferentes entre sí.

Para seleccionar el valor óptimo de k se aplica el método del codo. Este método calcula la suma de los errores cuadráticos dentro de cada clúster (SSE), con el fin de evaluar la variación interna de los grupos para diferentes valores k.

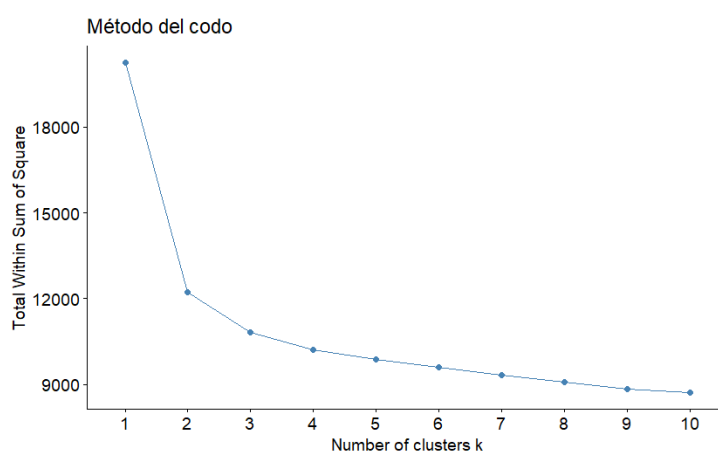


Figura 23: Método del codo

En base al gráfico (Figura 23), se observa una disminución notable en k igual a 2 indicando que con dos grupos se mejora la agrupación.

No obstante, se decide usar un valor k igual a 3. Esta decisión, se justifica, en primer lugar, porque la reducción de la variabilidad continúa siendo apreciable. Además, al ser un análisis social, la división en tres grupos será más fácil de interpretar permitiendo una interpretación más rica y matizada.

Desde un punto de vista analítico, la segmentación en tres clústeres proporciona una clasificación más diferenciada, facilitando la identificación de pensamientos intermedios que podrían quedar ocultas en una partición binaria.

Para mostrar los clústeres gráficamente se aplica un análisis de componentes principales como técnica de reducción de la dimensionalidad. Esta técnica transforma las variables originales en un nuevo conjunto de componentes principales, que son combinaciones lineales no correlacionadas entre sí. De esta manera representamos los clústeres y facilitamos su interpretación.

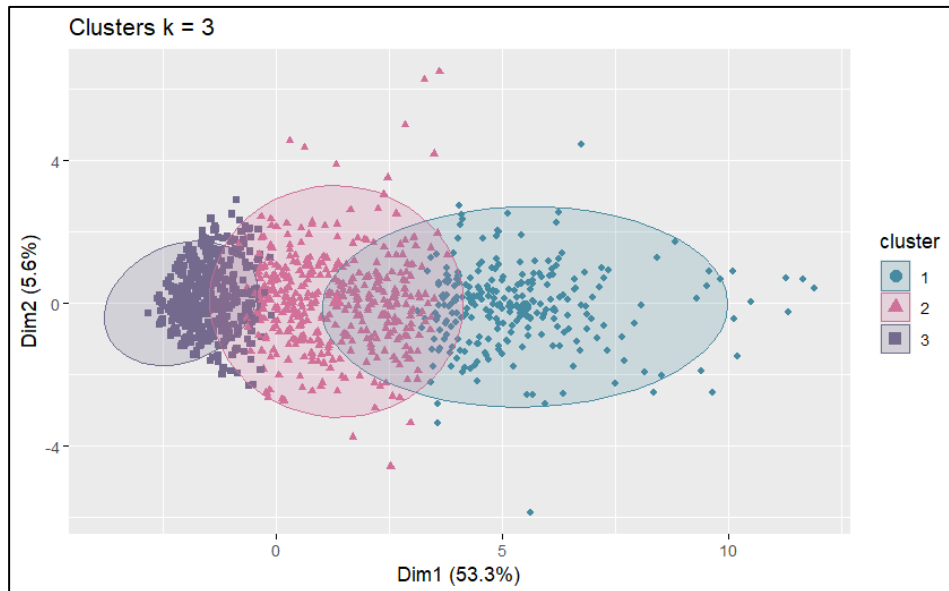


Figura 24: Visualización clústeres

En el gráfico (Figura 24) se representan los tres clústeres diferenciados por colores y formas. Se observa una diferenciación clara entre los clústeres, especialmente en el primer componente (Dim 1) que explica la mayor parte de varianza (53.3%). Aunque se observa cierto solapamiento entre el grupo 2 y 1, el clúster 1 se muestra más distanciados de los otros dos grupos, lo que indica un perfil más diferenciado en cuanto a la percepción de la masculinidad. A su vez, el primer clúster presenta mayor dispersión con respecto al tercero, que se muestra más compacto.

### 6.1. Caracterización de los clústeres

A continuación, se procede a la caracterización de cada clúster basándose en las variables que componen el análisis. Para lograrlo, se calculan las medias de respuesta de cada variable dentro de cada grupo, lo que permitirá encontrar patrones comunes de la percepción de la masculinidad en cada perfil, así como las diferencias entre ellos. Este análisis se realiza con el fin de identificar el significado social de los clústeres generados.

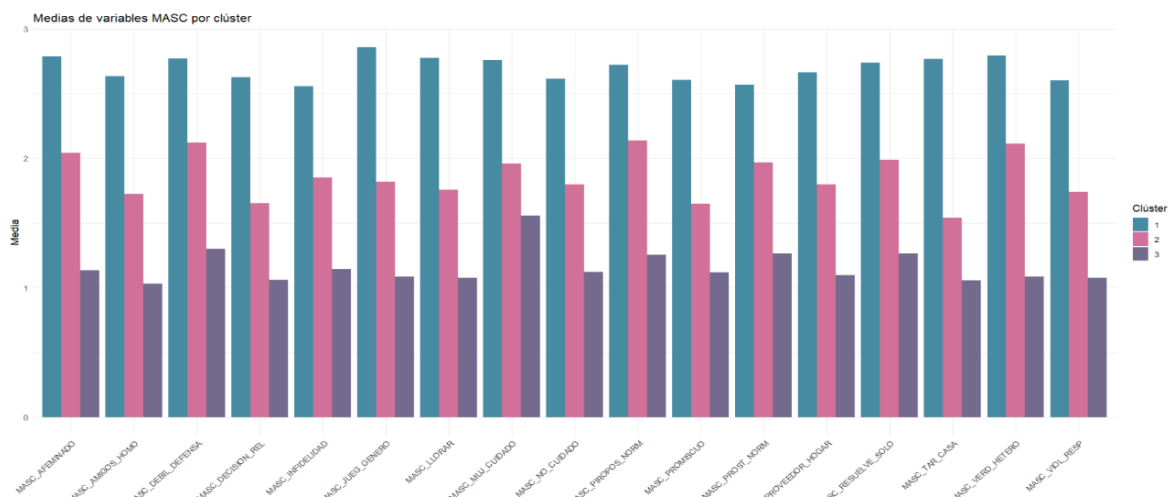


Figura 25: Media de respuestas de las variables MASC por clúster

Recordando que las variables sobre la percepción de la masculinidad toman valores del 1 al 4 (totalmente en desacuerdo – totalmente de acuerdo), el análisis de las medias permite identificar diferencias claras que existen entre los tres grupos.

En primer lugar, se observa (Figura 25) que las medias del primer clúster toman valores en torno a 2.9 en todas las variables, lo que indica alta aceptación de las afirmaciones asociadas a la masculinidad tradicional. Los individuos de este grupo tienden a estar de acuerdo con estereotipos que refuerzan los roles de género convencionales. Por ese motivo, se categoriza este perfil como “masculinidad tóxica” al representar una visión conservadora.

El clúster 2 muestra medias más bajas, en torno a 1.5, llegando a 2 en algunas variables. Estos individuos se alejan de los discursos más extremos de la masculinidad hegemónica, pero conservan creencias tradicionales. Este perfil intermedio se ha denominado “masculinidad neutra”, ya que refleja un punto de transición entre la aceptación en ciertos estereotipos tradicionales y la incorporación de actitudes más igualitarias.

Por último, el tercer clúster toma valores iguales o cercanos a 1, lo cual indica un rechazo sistemático a los estereotipos de la masculinidad tradicional. Los individuos de este grupo presentan desacuerdo con las normas sociales que perpetúan roles de género rígidos, lo que sugiere una mayor evolución social. Se categoriza este grupo como “masculinidad sana” al representar un modelo más flexible e inclusivo.

## 6.2. Perfilado de los clústeres

Se analiza la proporción de individuos existentes en cada clúster permitiendo obtener una primera aproximación a la percepción de la masculinidad.

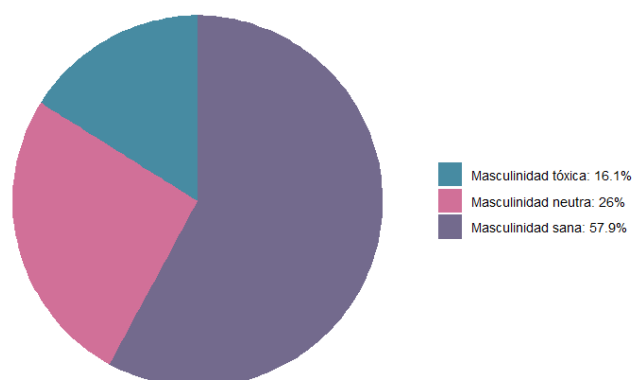


Figura 26: Distribución por clúster

Al observar la distribución de los clústeres (Figura 26), se identifica que el grupo con mayor proporción es el tercero, el correspondiente a la “masculinidad sana”. Aproximadamente el 58% de los individuos pertenecen a este grupo, algo positivo, pues supone una clara evolución social al reflejar mayor rechazo hacia los pensamientos tradicionales.

El 26% de los individuos encuestados pertenecen a la categoría “masculinidad neutra” donde mantienen esos pensamientos tradicionales, pero muestran cierto rechazo hacia esta tendencia en algunos ámbitos.

Por último, el 16% restante pertenece al grupo de “masculinidad tóxica”. Aun siendo la categoría minoritaria, su proporción no es despreciable. Esto implica que hay una parte significativa de la juventud que reproduce estos pensamientos tradicionales de la masculinidad.

Es interesante conocer la proporción de personas que pertenecen a cada clúster según su género para conocer si existe relación entre la percepción de la masculinidad y el género.

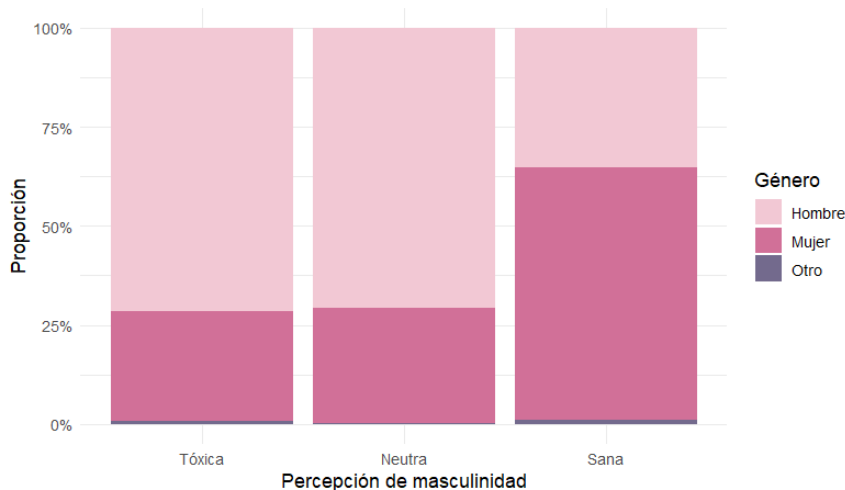


Figura 27: Distribución clústeres por género

Se observa (Figura 27) que aproximadamente el 70% de los individuos que pertenecen al primer y segundo clúster, “masculinidad tóxica” y “masculinidad neutra” son hombres. Por el contrario, en el caso del tercer clúster, “masculinidad sana”, la proporción de hombres desciende al 35%. Esto refleja que las percepciones de la masculinidad tradicionales son predominantes en los hombres, mientras que las mujeres y otras identidades tienden a alinearse con modelos de masculinidad igualitaria.

Para analizar si existe relación entre el género y los clústeres estadísticamente significativa, se realiza la prueba de independencia chi-cuadrado.

Tabla 12: Prueba independencia chi-cuadrado clúster y género

Chi-cuadrado	Grados de libertad	p-valor
211.21	4	< 2.2e-16

Se obtiene un p-valor pequeño, menor al umbral de significación 0.05, por lo que se rechaza la hipótesis nula de independencia y se concluye que existe una relación entre el género y la percepción de la masculinidad.

Esta relación indica mayor resistencia por parte de los hombres a asociarse con una masculinidad igualitaria y sana, mientras que las mujeres y otras identidades de género tienden a rechazar los pensamientos tradicionales.

Continuando con el perfilado de los clústeres, se analiza si existen relaciones estadísticamente significativas entre las variables sociodemográficas y los clústeres. Al ser variables categóricas (ORIENTACION, CLASE\_SOCIAL) se realiza la prueba chi-cuadrado, al igual que se ha hecho con la variable GENERO.

Se comienza realizando dicha prueba con la variable CLASE\_SOCIAL.

Tabla 13: Prueba independencia chi-cuadrado clase social y clúster

Chi-cuadrado	Grados de libertad	p-valor
91.878	8	< 2.2e-16

Se obtiene un p-valor inferior a 0.05, por lo que se rechaza la hipótesis nula y se concluye que existe una relación estadísticamente significativa. A continuación, se muestra la distribución de dicha variable por clúster.

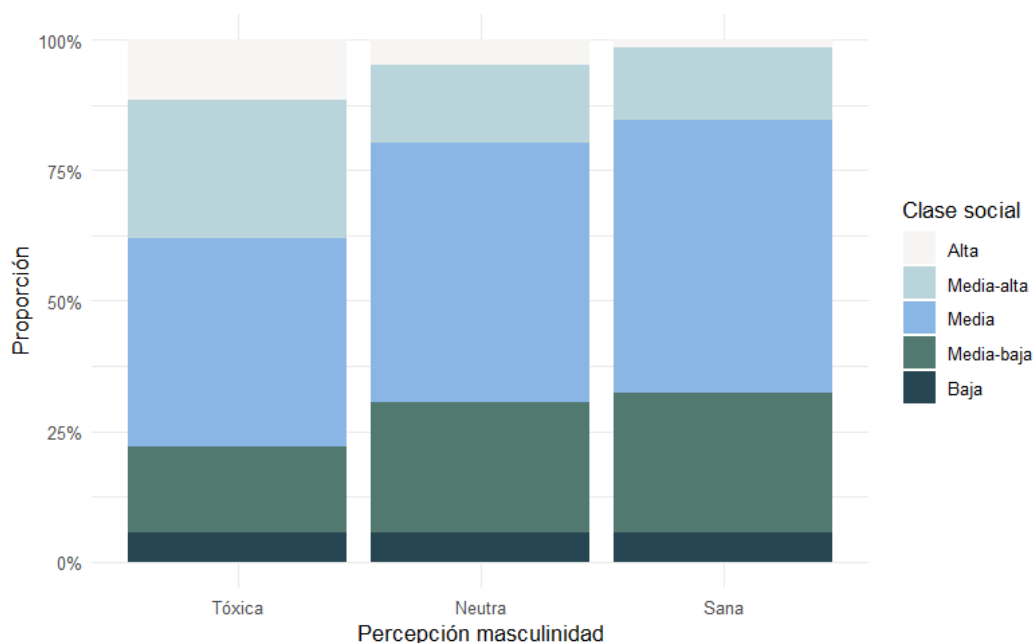


Figura 28: Distribución clase social por clúster

En el perfil tóxico se observa (Figura 28) una mayor proporción de clase alta y media alta en comparación con los otros dos grupos. Por otro lado, en los grupos neutro y sano, predominan las categorías medias y media baja.

Estos resultados sugieren que las percepciones tradicionales de la masculinidad son más comunes en clases sociales más altas.

A continuación, se realiza la prueba chi-cuadrado para la variable ORIENTACIÓN.

Tabla 14: Prueba independencia chi-cuadrado orientación y clúster

Chi-cuadrado	Grados de libertad	p-valor
104.82	6	< 2.2e-16

Al igual que en las dos variables anteriores, se rechaza la hipótesis nula de independencia y se concluye que existe relación estadísticamente significativa entre la variable ORIENTACIÓN y el clúster. Se procede a graficar la distribución de dicha variable por clúster.

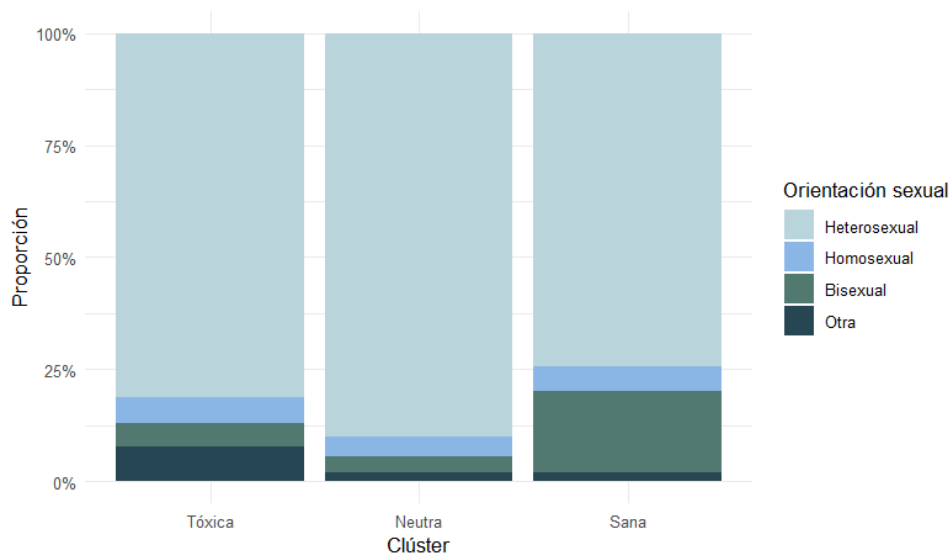


Figura 29: Distribución orientación sexual por clúster

Se observa (Figura 29) mayor diversidad sexual en el perfil sano que en el resto de los dos grupos, indicando que los perfiles de masculinidad más igualitarios tienden a incluir mayor representación de orientaciones sexuales no heteronormativas.

Cabe destacar, que se observa una mayor proporción de la categoría “otra” en el grupo tóxico frente a los dos grupos restantes. Puede parecer contradictorio a primera vista, sin embargo, esto podría explicarse por el carácter residual de dicha variable, que agrupa identidades diversas no siempre con una visión progresista. Asimismo, se observa que en grupos sexualmente diversos pueden mantenerse posturas tradicionales.

Se concluye este bloque realizando el test Kruskal-Wallis que contrasta la hipótesis nula de que las medianas de todos los grupos son iguales, con la variable EDAD para identificar si existe relación estadísticamente significativa entre dicha variable y la percepción de la masculinidad.

Tabla 15: Test Kruskal-Wallis de edad y clúster

Kruskal-Wallis	Grados de libertad	p-valor
104.82	2	0.1439

Se obtiene un p-valor igual a 0.14, y, por tanto, superior al umbral de significación 0.05. Por lo que no podemos rechazar la hipótesis nula y se concluye que no se encuentran diferencias estadísticamente significativas de la edad entre los distintos clústeres.

Tras analizar las relaciones de las variables socio demográficas y la percepción de la masculinidad, se procede a examinar la distribución de los clústeres en función de variables relacionadas con la salud mental y el bienestar personal.

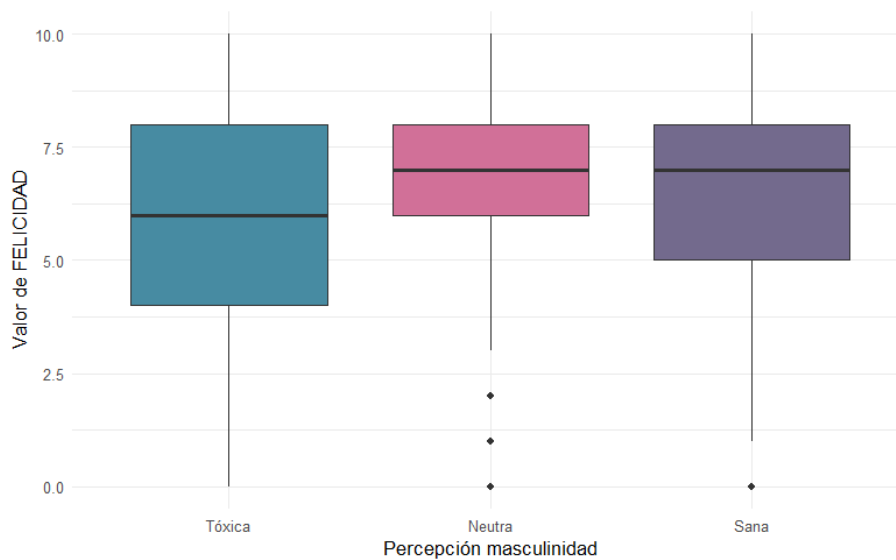


Figura 30: Distribución felicidad por clúster

En primer lugar, se observa la distribución de la variable FELICIDAD por clúster (Figura 30). Las medianas en los clústeres neutro y sano son ligeramente superiores al clúster tóxico. Además, la varianza en este último grupo es mayor, lo que refleja una mayor variabilidad emocional en este perfil, así como mostrar niveles de bienestar menos elevados.

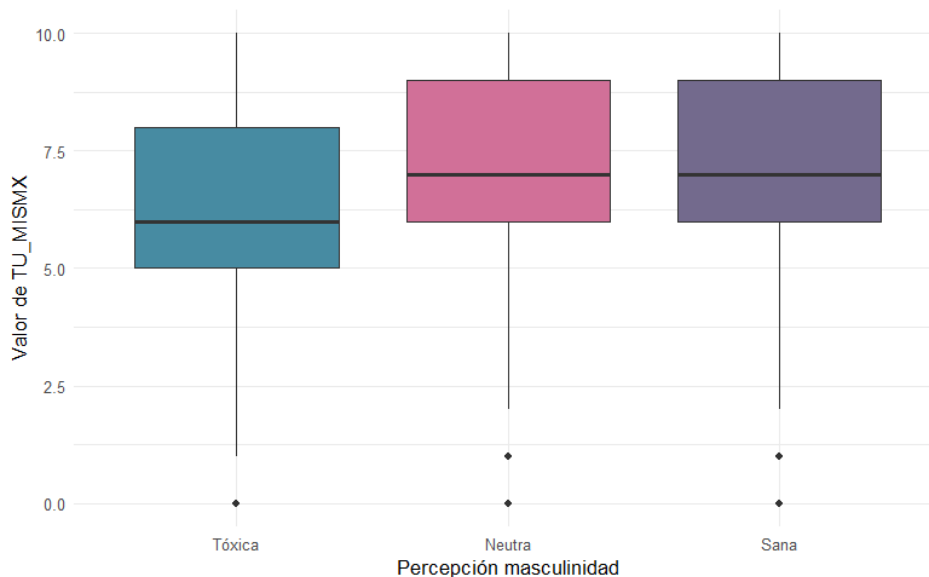


Figura 31: Distribución de sentirse una/o misma/o por clúster

En cuanto a la variable TU\_MISM (Figura 31), el grupo tóxico presenta una mediana inferior frente a los otros dos grupos restantes. Esto sugiere que las personas que perciben la masculinidad más tradicional tienen mayor dificultad por sentirse ellas mismas. Podría estar relacionado con sentir una mayor presión social por ser y/o actuar de determinada manera.

Por último, se observa la distribución de salud mental de los individuos por clúster.

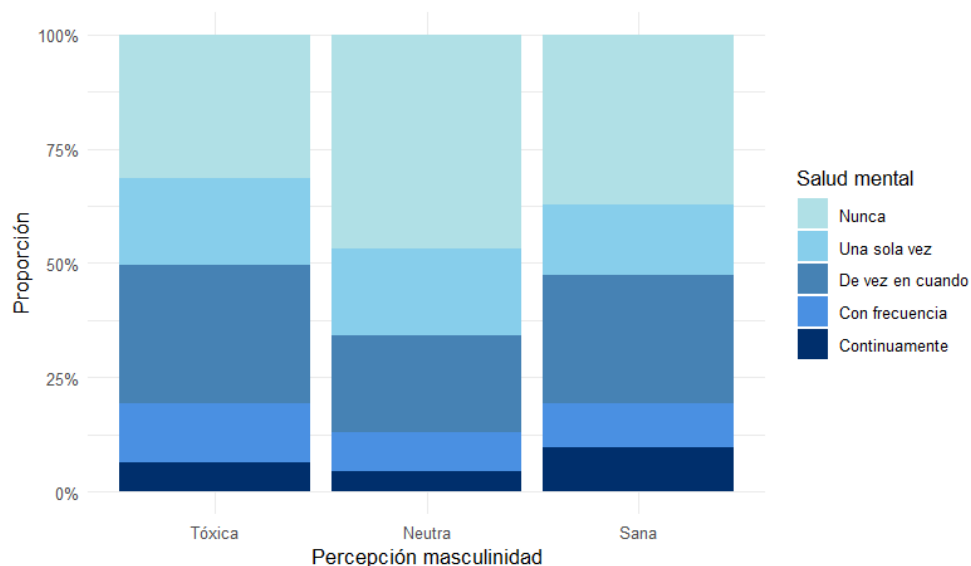


Figura 32: Distribución de salud mental por clúster

Se aprecia (Figura 32) que el perfil sano presenta una mayor frecuencia en la categoría “continuamente”. Sin embargo, en las categorías “con frecuencia” y “de vez en cuando” muestra proporciones similares al primer grupo (tóxico). Por otro lado, el grupo neutro tiene mayor frecuencia en la categoría “nunca” con respecto a los otros dos grupos. A simple vista, no se observan diferencias marcadas entre los grupos sano y tóxico, lo que sugiere ligera homogeneidad en las respuestas.

Con el fin de analizar si la masculinidad tradicional afecta a la salud mental de los hombres, se realiza un análisis comparativo entre hombres y mujeres. Esta distinción permite evaluar si la percepción de la masculinidad tradicional se asocia de forma diferente al bienestar emocional según el género.

Se aplica la prueba de independencia chi-cuadrado para cada uno de los grupos.

Tabla 16: Prueba independencia chi-cuadrado salud mental mujeres y clúster

Chi-cuadrado (mujeres)	Grados de libertad	p-valor
16.315	8	0.03808

Tabla 17: Prueba independencia chi-cuadrado salud mental hombres y clúster

Chi-cuadrado (hombres)	Grados de libertad	p-valor
19.092	8	0.00143

Aunque en ambas pruebas se observa que existe relación estadísticamente significativa al nivel del 0.05, se identifica mayor fuerza estadística en el grupo hombres. Lo que podría suponer que la percepción de la masculinidad puede afectar como factor de riesgo.

A continuación, se muestra la distribución de la salud mental por clúster dentro del subconjunto de hombres, para identificar esas posibles diferencias.

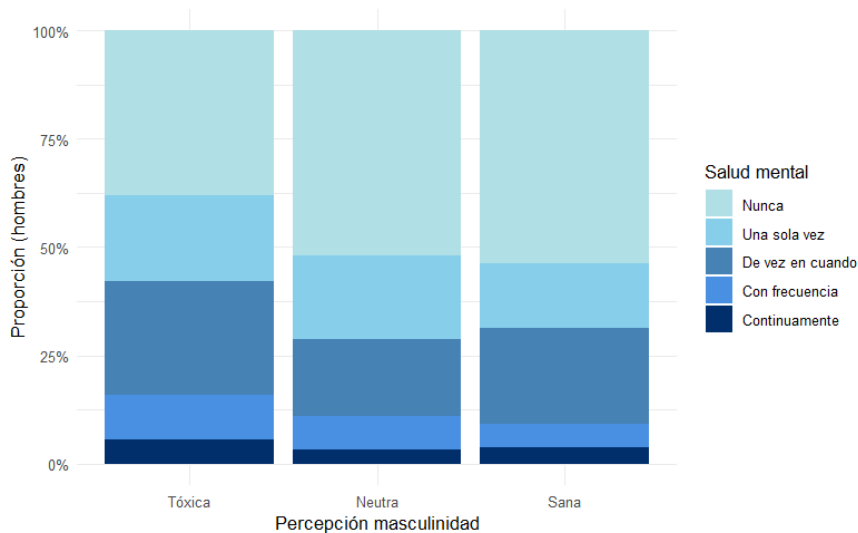


Figura 33: Distribución salud mental (hombres) por clúster

Se observa (Figura 33) que en este caso el perfil tóxico muestra mayor proporción en las categorías de malestar psicológico (“continuamente”, “con frecuencia” y “de vez en cuando”). Por otra parte, los hombres con una percepción sana o neutra muestran mayor presencia en la categoría “nunca”. Esto podría indicar, que una percepción más tradicional hacia la masculinidad puede estar asociado a un mayor sufrimiento emocional entre los hombres, posiblemente como consecuencia de las limitaciones que imponen los estereotipos tradicionales.

Para concluir el proceso de perfilado de los clústeres, se realiza un análisis para comparar las distintas variables relacionadas con la percepción de presiones sociales con los tres perfiles de masculinidad. El objetivo es identificar si existen diferencias entre los clústeres o qué presiones son más intensas o relevantes para cada perfil. Con este fin, se realiza un análisis del nivel medio de presión en función del clúster.

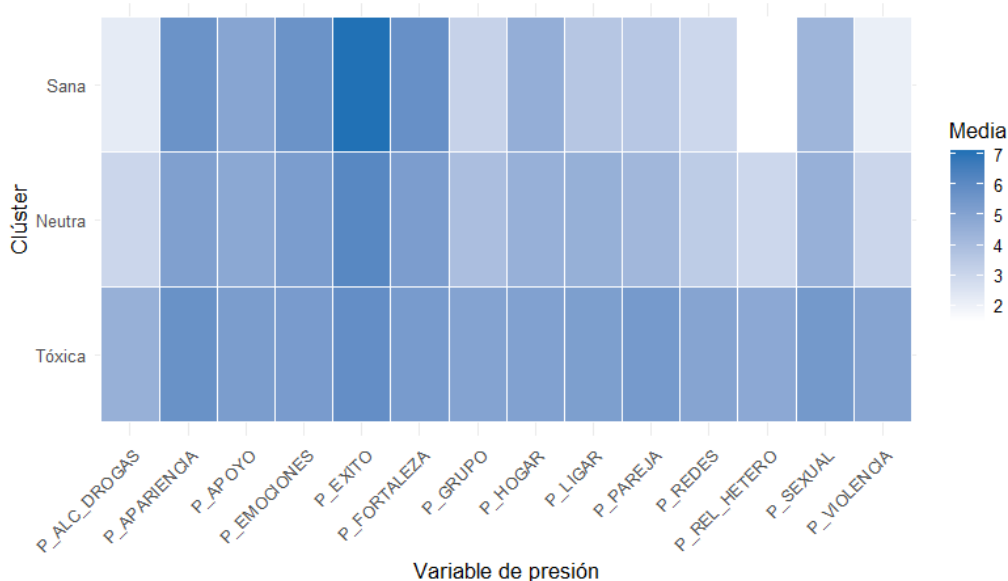


Figura 34: Nivel medio de presión social percibida según clúster

En general, el clúster de masculinidad tóxica refleja medias más homogéneas y elevadas en la mayoría de las variables. Una excepción destacable es P\_EXITO que se muestra más elevada en el grupo sano.

El clúster de masculinidad neutra ocupa una posición intermedia con niveles de presión moderados.

Llaman especialmente la atención las diferencias de las variables P\_ALC\_DROGAS, P\_GRUPO, P\_VIOLENCIA y P\_REL\_HETERO, donde la distancia entre los grupos 1 y 3 es superior. El grupo sano refleja medias pequeñas, mientras que el grupo tóxico muestra niveles de presión más elevados. Estas variables hacen referencia a presiones por consumir alcohol o drogas, encajar en un grupo, validación de conductas violentas y tener amistades únicamente heterosexuales. La diferencia entre los clústeres sugiere que los individuos con una visión más tradicional sienten mayor necesidad de ajustarse a estos mandatos sociales, lo que refleja que una masculinidad más tradicional se vincula a dinámicas de riesgo o exclusión. Por el contrario, los individuos que pertenecen al perfil sano no están condicionados por estas exigencias externas.

## 7. Análisis estadístico

Con el fin de entender e identificar las relaciones socioculturales y conductuales de los individuos con la percepción de la masculinidad, se procede a realizar modelos explicativos. La variable objetivo de dichos modelos será la creada previamente en el análisis clúster, donde se identificaron tres perfiles diferenciados: masculinidad tóxica, neutra y sana.

Se realizan modelos de regresión logística multinomial, técnica útil cuando la variable dependiente presenta más de dos categorías. Se aplican tres métodos diferentes de selección de variables, Stepwise, Forward y Backward. Cada uno de ellos, ha sido evaluado utilizando dos criterios de penalización, AIC (Akaike Information Criterion) y BIC (Bayesian Information Criterion). Dichos criterios evalúan el equilibrio entre el ajuste del modelo y su complejidad, penalizando el exceso de variables innecesarias. En total, se generan seis modelos distintos, mediante combinaciones de métodos de selección y criterio de ajuste.

Además, se ha incluido un modelo de árbol de clasificación, ya que la variable objetivo es de tipo categórico. Este modelo, permite representar visualmente la clasificación de los distintos perfiles y obtener una interpretación jerárquica de las variables que mejor explican las diferencias entre las categorías.

### 7.1. Monte Carlo Cross-Validation

Para garantizar la robustez de los modelos, se ha implementado una estrategia de validación conocida como Monte Carlo Cross Validation (MCCV). Antes de detallar en profundidad el proceso, se muestra un mapa de flujo (Figura 35) que permite observar de forma ordenada las distintas etapas de esta estrategia.

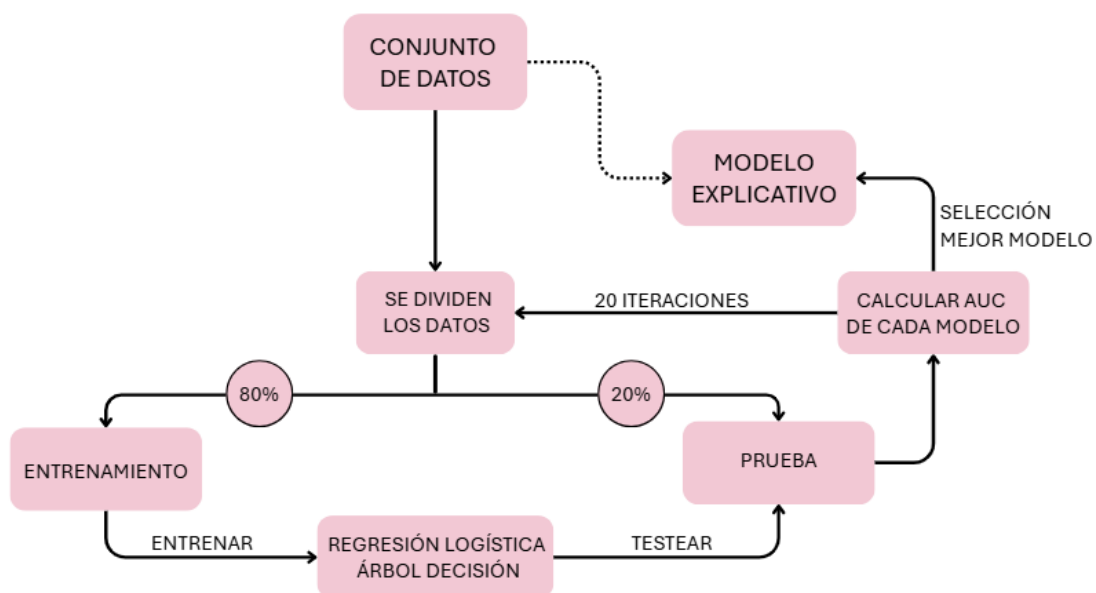


Figura 35: Mapa de flujo de Monte Carlo Cross-Validation

Este procedimiento consiste en realizar 20 particiones aleatorias independientes del conjunto de datos. En cada iteración, los datos se dividen en dos subconjuntos, el 80% para el entrenamiento de los modelos y el 20% restante para su validación. Esta técnica permite estimar de forma más fiable el rendimiento de los modelos.

La calidad de cada modelo se evalúa mediante el área bajo la curva ROC (AUC). Esta medida captura la capacidad del modelo en distinguir correctamente entre las diferentes categorías de masculinidad. Se registran los valores AUC obtenidos en cada una de las particiones, y se representa los resultados en boxplots.

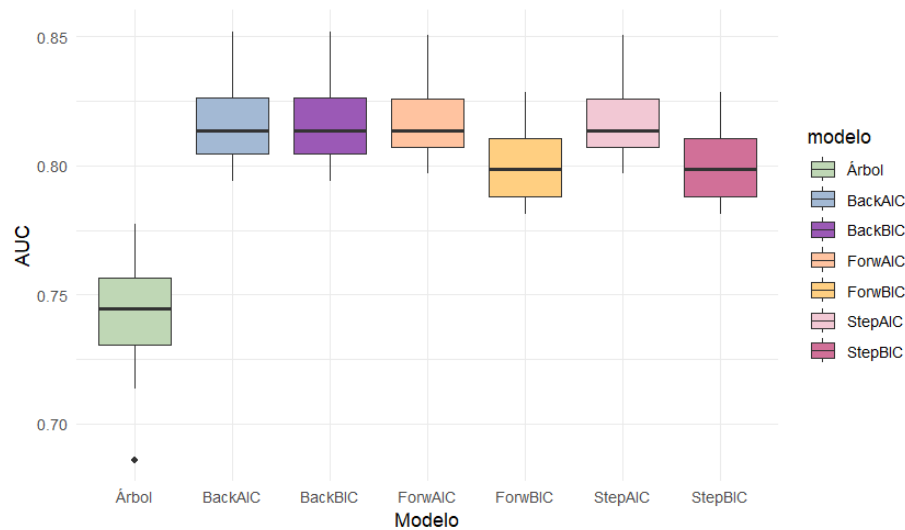


Figura 36: Comparación AUC entre modelos con boxplot

Se observa (Figura 36) que existen igualdad de métricas para cada par de modelos (BackAIC y BackBIC, ForwAIC y StepAIC, ForwBIC y StepBIC). Esto se debe a que, aunque se han aplicado distintos métodos de selección y penalización, el conjunto final ha resultado ser el mismo y, por tanto, presentan las mismas medidas AUC. Para simplificar su análisis comparativo, se agrupan del siguiente modo:

- Modelo 1: BackAIC y BackBIC.
- Modelo 2: ForwAIC y StepAIC.
- Modelo 3: ForwBIC y StepBIC.

En primer lugar, se destaca que los modelos de regresión presentan una calidad predictiva superior que la que muestra el árbol de decisión. Esta diferencia puede deberse a la naturaleza de ambos modelos. Mientras que la regresión logística asume relaciones lineales entre las variables independientes y la objetivo, el árbol de decisión captura relaciones no lineales y jerárquicas. Por lo tanto, la estructura de los modelos lineales se ajusta mejor a las relaciones presentes en los datos.

No obstante, se mantiene el árbol de decisión en el análisis por su valor complementario, ya que es un modelo fácilmente interpretable y puede mostrar relaciones que el modelo de regresión no capture. Además, presenta valores AUC que oscilan entre 0.73 y 0.76 aproximadamente, lo que implica que logra clasificar correctamente una proporción considerable de los casos.

En segundo lugar, se observan diferencias relativamente pequeñas entre los modelos de regresión logística multinomial. Todos ellos muestran valores AUC en torno a 0.83, es decir, tienen una alta capacidad predictiva. Podemos observar que el modelo 3 muestra peor rendimiento, con una mediana de AUC más baja (0.8) y una mayor dispersión. Los modelos

1 y 2 presentan una mediana AUC prácticamente igual, sin embargo, el modelo 2 tiene una varianza ligeramente menor, lo que sugiere mayor consistencia y robustez. Por ese motivo es el más adecuado en cuanto a calidad.

Para completar la evaluación, se analizará la complejidad de los modelos, para valorar el equilibrio entre capacidad explicativa y complejidad. Para ello se ha realizado un seguimiento de las variables incluidas en cada uno de los modelos a lo largo de las 20 particiones realizadas. Para evaluarlo de manera objetiva, se aplica un enfoque basado en la frecuencia de selección de variables. En concreto, se mostrarán únicamente aquellas variables que han sido seleccionadas en al menos 15 particiones (75%), para evitar mostrar aquellas que se incluyeron en los modelos debido a una partición específica favorable y mostrar aquellas variables que su relación con la variable objetivo no es producto del azar muestral.

Tabla 18: Variables incluidas por modelo de regresión

Variables	BackAIC/BackBIC (modelo 1)	ForwAIC/StepAIC (modelo 2)	ForwBIC/StepBIC (modelo 3)
AVERGONZAR	✓	✓	✓
BURLA	✓	✓	
FEMINISTA	✓	✓	✓
GENERO	✓	✓	✓
HABLAR_AMIGXS	✓	✓	
H_VERDAD_ENTORNO	✓	✓	
M_VERDAD_ENTORNO	✓	✓	
IDEOLOGIA	✓	✓	
ORIENTACION	✓	✓	✓
P_EXITO	✓	✓	
P_REL_HETERO	✓	✓	✓
P_VIOLENCIA	✓	✓	
RELIGIOSIDAD	✓	✓	
P_ALC_DROGAS	✓	✓	
VIOLENCIA	✓	✓	✓
TU_MISMIX	✓	✓	
P_FORTALEZA	✓	✓	
P_EMOCIONES	✓	✓	
P_PAREJA		✓	

El resto de las variables que no se encuentran en la tabla (Tabla 18) han sido incluidas en los modelos con una frecuencia menor que 15 a lo largo de las 20 particiones del conjunto de datos. Dichas variables son P\_GRUPO, FELICIDAD, P\_HOGAR, P\_LIGAR, CLASE\_SOCIAL, EDAD, INDICE\_RIESGO, P\_REDES y P\_SEXUAL. Cabe destacar que estas variables fueron incluidas con una frecuencia igual o menor a seis, lo que sugiere que, fue debido a particiones específicas favorables.

En cuanto a la comparación entre modelos, se observa que los modelos 1 y 2 seleccionan casi el mismo número de variables de forma general, sin embargo, se observan ligeras diferencias en algunas particiones. Esta variabilidad provoca que el modelo 1 muestre una

mayor dispersión en los valores AUC, como se muestra en el gráfico anterior (Figura 36). Esto puede estar asociado a una menor estabilidad o un leve sobreajuste.

En base a estos resultados, se selecciona el modelo 2 como el mejor modelo explicativo. Este modelo presenta mayor calidad predictiva, alcanzando valores de AUC más altos (en torno a 0.83) y una menor varianza en comparación con los dos modelos de regresión restantes. En esta decisión, no se prioriza el principio de parsimonia, que recomienda seleccionar el modelo más simple entre opciones de calidad similar, ya que el objetivo del trabajo es explicativo y no predictivo. Por tanto, se opta por el modelo que mejor permite identificar las relaciones existentes entre las variables y la percepción de la masculinidad.

## 7.2. Modelo regresión logística multinomial

Se ha seleccionado el modelo 2 como el mejor al presentar la mayor calidad predictiva y menor variabilidad. Dicho modelo se ha obtenido tanto por el método Stepwise con AIC como por el método Forward con AIC, ya que ambos procesos han resultado dar el mismo conjunto final de variables.

Para la interpretación del modelo, se ha optado por presentar los resultados obtenidos mediante el proceso StepAIC, y el modelo ha sido reentrenado con el conjunto de datos total.

Dado que los coeficientes de regresión logística multinomial se expresan en términos de log-odds, se procede a transformarlos en *odds ratio* (OR). Este valor representa el cambio relativo de probabilidad de pertenecer a una categoría determinada frente a la de referencia (en este caso, “masculinidad tóxica”).

En la siguiente tabla se presentan los OR para las categorías neutra y sana, lo que permite identificar como cambia la probabilidad de pertenecer a cada uno de estos perfiles en función de las variables explicativas del modelo.

Tabla 19: OR del modelo de regresión logística multinomial con categoría "tóxica" como referencia

PARÁMETRO	NEUTRA	SANA
AVERGONZAR	0.6583	0.4971
FEMINISTA1	1.1605	4.0392
FEMINISTA2	1.0597	1.9436
GENERO2	1.0976	2.8684
GENERO3	0.8076	2.4841
P_REL_HETERO	0.9181	0.8057
VIOLENCIA	0.7509	0.6725
RELIGIOSIDAD	0.9594	0.8911
P_EXITO	1.0762	1.1642
ORIENTACION2	1.4022	2.8984
ORIENTACION3	0.7771	2.2214
ORIENTACION4	0.2789	0.3495
BURLA	0.8511	0.655
HABLAR_AMIGXS2	1.6993	1.7497
HABLAR_AMIGXS3	1.5317	2.7153
HABLAR_AMIGXS4	1.2137	1.2407
P_ALC_DROGAS	0.9611	0.9048
IDEOLOGIA	1.0242	0.8991
TU_MISMIX	1.1083	1.1631
P_VIOLENCIA	0.8958	0.9008
H_VERDAD_ENTORNO1	1.1233	0.4927
M_VERDAD_ENTORNO1	0.6284	1.1922
P_FORTALEZA	1.0753	1.1179
P_EMOCIONES	1.0776	1.0858
P_PAREJA	0.948	0.9232

Se muestran en rosa los parámetros que son estadísticamente significativos al 5%, es decir, tenemos evidencias de que estos efectos sean distintos de 0, por lo tanto, se interpretan solo aquellos que están marcados.

- AVERGONZAR: a medida que aumenta la frecuencia de que el individuo avergüence a una persona se reducen las posibilidades al 35% de pertenecer a la clase neutra frente a la tóxica. En el caso de la clase sana, la probabilidad disminuye en un 51% frente al perfil tóxico. Lo que indica que esta conducta está más asociada hacia un pensamiento más tradicional o tóxico.
- FEMINISTA: aquellos individuos que se consideran feministas tienen 4.03 veces más probabilidad de pertenecer a la clase sana que a la tóxica frente a no considerar serlo. Asimismo, quienes responden NS/NC a esta pregunta será 94% más probable pertenecer a la clase sana respecto a aquellas personas que rechazan el feminismo. Esto sugiere que el feminismo es un factor fuertemente asociado a las percepciones sobre la masculinidad más alejadas del pensamiento tradicional.

- **GENERO:** las mujeres tienen casi 3 veces más probabilidad de pertenecer a la clase sana que a la tóxica frente a los hombres. Esto muestra que las mujeres presentan mayor rechazo hacia las percepciones de la masculinidad tradicionales.
- **P\_REL\_HETERO:** cuanto mayor es la presión de tener únicamente amistades con personas heterosexuales, la probabilidad de pertenecer al perfil de masculinidad neutra disminuye un 9% frente a la tóxica. En el caso del perfil sano, disminuye un 20% la probabilidad de pertenecer a dicha clase frente a la tóxica.
- **VIOLENCIA:** A medida que aumenta la frecuencia con la que un individuo ejerce violencia, la probabilidad de pertenecer al grupo de masculinidad neutra se reduce en un 25% frente al perfil tóxico. En el caso de la masculinidad sana, la probabilidad disminuye aún más, en un 33%, lo que refuerza la asociación entre conductas violentas y formas tradicionales de masculinidad.
- **RELIGIOSIDAD:** un mayor grado de religiosidad disminuirá un 11% la probabilidad de pertenecer a la clase sana frente a pertenecer a la tóxica. No obstante, cabe señalar que esta variable presentó un ligero error de imputación durante el proceso de tratamiento de datos, por lo que sus resultados deben interpretarse con cierta precaución.
- **P\_EXITO:** cuanto mayor es la presión de tener éxito en el trabajo aumenta un 16% la probabilidad de pertenecer a la clase sana frente a la tóxica.
- **ORIENTACION:** Las personas homosexuales tienen casi 3 veces más probabilidad de pertenecer a la clase sana que a la tóxica frente a las personas heterosexuales. Además, las personas bisexuales, tienen 2.22 veces más probabilidad de pertenecer a la clase sana que a la tóxica, frente a las personas heterosexuales. Esto sugiere que las orientaciones no heteronormativas están más alejadas del pensamiento tradicional.
- **BURLA:** cuanto mayor es la frecuencia de que un individuo se burle de una persona se reduce las posibilidades de pertenecer a la clase neutra al 20%, y al 35% en el caso de pertenecer a la clase sana, frente a la tóxica. Este patrón refuerza la asociación entre actitudes humillantes y percepción de masculinidad tradicional.
- **HABLAR\_AMIGXS:** tener tanto amigos como amigas con los que hablar de temas personales aumenta un 2.7 veces más probable pertenecer a la clase sana frente a la tóxica que no contar con nadie con quien hablar. Estos valores destacan que tener apoyo emocional de amistades está vinculado a un rechazo por la masculinidad tradicional.
- **P\_ALC\_DROGAS:** a medida que aumenta la presión por consumir alcohol o drogas disminuye un 10% la probabilidad de pertenecer a la clase sana frente a la tóxica. Lo que evidencia que estas prácticas de riesgo están más asociadas a perfiles tradicionales o tóxicos.
- **IDEOLOGÍA:** cuánto más se identifica un individuo con políticas de derechas, menor será la probabilidad, un 11%, de pertenecer a la clase sana frente a la tóxica.
- **TU\_MISMIX:** a medida que una persona siente mayor libertad por mostrarse ella misma incrementa un 16% la probabilidad de pertenecer a la clase sana frente a la tóxica. Además, esta probabilidad aumenta un 10% de pertenecer a la clase neutra frente a la tóxica.
- **P\_VIOLENCIA:** a medida que aumenta la presión por usar la violencia para defender ideas disminuye un 10% aproximadamente la probabilidad de pertenecer a la clase neutra o sana frente a la tóxica.

- M\_VERDAD\_ENTORNO: los individuos que escuchan en su entorno que la “mujer de verdad” debe comportarse de una manera determinada tienen un 40% menos probabilidad de pertenecer a la clase neutra frente a la tóxica.
- H\_VERDAD\_ENTORNO: los individuos que escuchan en su entorno que un “hombre de verdad” debe comportarse de una manera determinada tienen 51% menos probabilidad de pertenecer a la clase sana que a la tóxica.
- P\_FORTALEZA: a medida que aumenta la presión por mostrar fortaleza en momentos de nerviosismo o vulnerabilidad aumenta un 12% la probabilidad de pertenecer a la clase sana frente a la tóxica. Lo que podría indicar que el concepto de fortaleza se asocia a pensamientos tradicionales.
- P\_EMOCIONES: a medida que aumenta la presión por ocultar la tristeza o ansiedad aumenta un 8% la probabilidad de pertenecer tanto a la clase neutra como a la sana frente a la tóxica.
- P\_PAREJA: a medida que aumenta la presión por tener pareja estable disminuye un 8% la probabilidad de pertenecer la clase sana frente a la tóxica, sugiriendo que la normatividad relacional, influye ligeramente en el modelo de masculinidad adoptado.

### 7.3. Árbol de decisión

Con el fin de complementar el análisis, se realiza un árbol de decisión. Este proceso tiene como objetivo contrastar los resultados e identificar posibles relaciones no lineales en los datos. Dado que la variable objetivo es de tipo categórica, se procede a realizar un árbol de clasificación.

Se entrena el modelo con el 100% de la muestra, lo que permite utilizar toda la información contenida en los datos para nuestro objetivo explicativo.

En primer lugar, se crea un primer árbol sin restricciones de complejidad para capturar toda la estructura del conjunto de datos, dicho árbol se define como maximal. Se ha definido un valor mínimo de hoja pequeño, siendo el 1.5% de los datos totales, lo que permite detectar subdivisiones relevantes.



tamaños de hoja posibles, por ejemplo, pasa de 4 a 7 hojas. Esto es frecuente que ocurra en los árboles de clasificación, ya que el número de observaciones clasificadas incorrectamente se mantiene constante en estos casos.

Se aprecia una línea de puntos en el gráfico (Figura 38), todos los árboles que se encuentran debajo de este umbral no son significativamente distintos entre sí desde un punto de vista predictivo. Se observa que el primer árbol por debajo de la línea tiene un número de hojas igual a 7. También, se aprecia que el árbol con el menor valor del error se trata del árbol con un número de hojas igual a 9. Por lo tanto, cualquiera de estas opciones sería adecuada.

Para sacar conclusiones más robustas, se muestra el gráfico del AUC dependiendo del número de hojas.

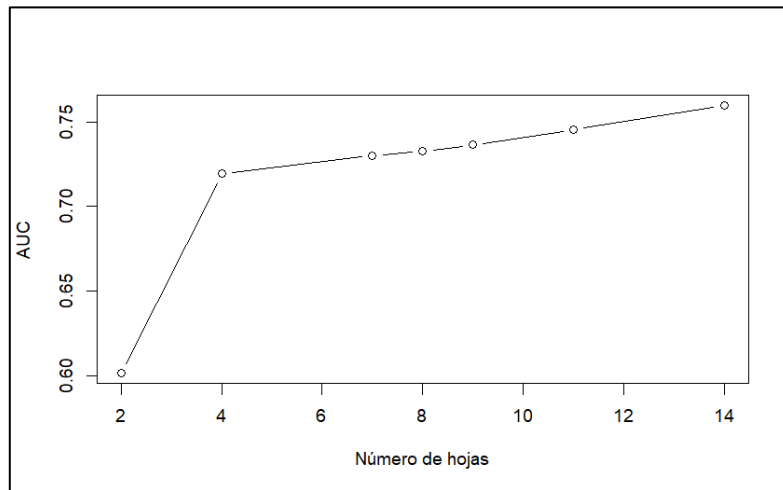


Figura 39: Valor AUC por número hoja

Se detecta (Figura 39) que tanto con 7 como con 9 hojas se obtienen valores de AUC aproximadamente igual a 0.73 siendo el segundo árbol, levemente superior al primero. Haciendo uso del principio de parsimonia y observando que la variación del valor AUC de 7 a 9 hojas no es significativa, seleccionaremos finalmente como valor óptimo el árbol con 7 hojas, para obtener un árbol que presente mayor facilidad de interpretación. Tras la selección del valor óptimo, representamos el árbol podado.

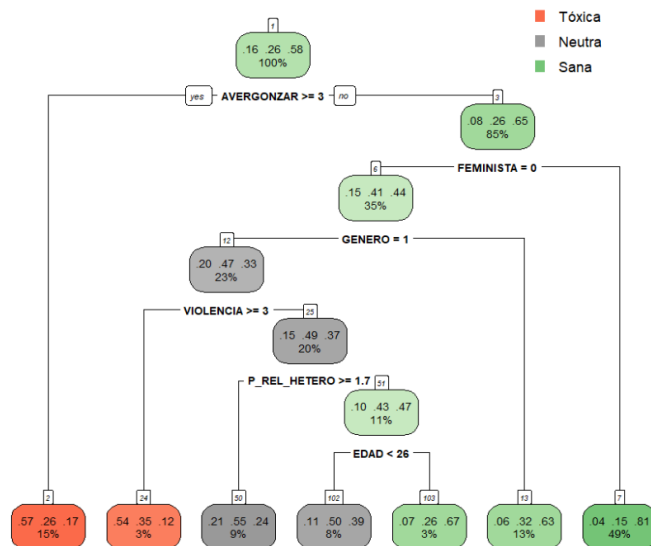


Figura 40: Árbol podado

El árbol podado tiene una profundidad igual a 6 con 7 hojas. Se observa en cada nodo la proporción de individuos existentes de cada categoría y el porcentaje del conjunto de datos que se encuentra en cada nodo. El árbol se representa mediante colores dependiendo de la categoría mayoritaria de cada nodo, siendo un color más intenso en aquellos nodos donde la proporción de la categoría mayoritaria sea muy superior al resto.

Se observa que el árbol comparte la mayoría de las variables con el modelo de regresión logística creado anteriormente, a excepción de la variable EDAD, lo que puede indicarnos que existe una relación no lineal entre esta variable y la percepción de la masculinidad.

### 7.3.1. Interpretación del árbol de decisión

A continuación se explicarán las características de cada hoja del árbol resultante y se mencionará la proporción de la categoría mayoritaria en cada nodo.

Hojas clasificadas como masculinidad tóxica:

- NODO 2: Este nodo representa el 15% del total de los datos. El 57% de los individuos de este nodo pertenecen al perfil de masculinidad tóxica, el 26% pertenece al grupo neutro y el 17% restante pertenece a la clase sana. Corresponde a aquellos individuos que avergüenzan a otras personas con mucha frecuencia, lo que sugiere fuerte asociación entre los pensamientos tradicionales y actos humillantes.
- NODO 24: Se agrupa el 3% del conjunto total de los datos. El 54% de ellos pertenecen al primer grupo (tóxica), siendo el 35% pertenecientes al perfil neutro y el 12% restante al grupo de masculinidad sana. Se encuentran aquellos individuos hombres que no avergüenzan a otras personas o lo realizan con poca frecuencia, no se consideran feministas y ejercen violencia con mucha frecuencia. De nuevo se observa actos humillantes y agresivos que se relacionan con el grupo tóxico.

Hojas clasificadas como masculinidad neutra:

- NODO 50: Agrupa el 9% del total del conjunto de datos. El 55% de los individuos son del perfil neutro, siendo la proporción de los grupos restantes similar (24%

pertenecen al grupo sano y el 21% restante al grupo tóxico). Se encuentran aquellos hombres que no avergüenzan a otras personas o lo realizan con poca frecuencia, no se consideran feministas, no ejercen violencia o la realizan con poca frecuencia y muestran presión por relacionarse con personas únicamente heterosexuales.

- NODO 102: Se encuentra el 8% del total de los datos del cual el 50% pertenece al grupo neutro, siendo el 39% perteneciente al grupo sano y el 11% restante perteneciente al grupo tóxico. Agrupa a aquellos hombres que no avergüenzan a otras personas o lo realizan con poca frecuencia, no se consideran feministas, no ejercen violencia o la realizan con poca frecuencia, no muestran presión por relacionarse con personas únicamente heterosexuales y tienen menos de 26 años.

Se puede observar que el grupo neutro comparte características con los dos grupos restantes. Se encuentran hombres que no presentan actos agresivos, pero mantienen ciertos pensamientos heteronormativos, lo que podría indicar resistencia a masculinidades sanas. Además, se observa que en el nodo 51 se realiza una partición con la variable edad. Esto nos indica que los más jóvenes (menores de 26), presentan cierto rechazo hacia las masculinidades más equitativas con respecto a los mayores de 26.

Hojas clasificadas como masculinidad sana:

- NODO 103: Este nodo representa el 3% de los datos totales. El 67% pertenece al perfil sano, el 26% al grupo neutro y tan solo el 7% restante pertenece al grupo tóxico. Este nodo comparte todas características con el nodo 102, salvo que en este nodo se encuentran los jóvenes con edad igual o superior que 26. Contrario a lo que podría esperarse, implica un retroceso, ya que los más jóvenes presentan un mayor rechazo hacia la masculinidad sana frente a los mayores de 26.
- NODO 13: Se encuentra el 13% del total, de los cuales, el 63% pertenece al grupo sano, siendo el 36% pertenecientes al grupo neutro y el 6% restante pertenecientes al perfil tóxico. En este nodo se encuentran aquellas personas que no avergüenzan a otras personas o lo realizan con poca frecuencia, se identifican con género distinto a hombre y no se consideran feministas. Aún sin una identificación explícita del feminismo, las mujeres y personas de otro género distinto al hombre se asocian con pensamientos sobre masculinidad más igualitarios.
- NODO 7: Se encuentra casi la mitad del total de los datos (49%). El 81% de los individuos pertenecen al tercer grupo (sano), siendo el 15% y el 4% restantes pertenecientes al grupo neutro y tóxico respectivamente. En este nodo se encuentran aquellos individuos que no avergüenzan a otras personas o lo realizan con poca frecuencia, además se consideran feministas o respondieron con NS/NC a dicha pregunta.

Se observa que las mujeres y personas que se identifican con otro género tienden a adoptar pensamientos más igualitarios. El nodo 7, representa casi la mitad de la muestra, evidencia que la identificación con el feminismo o ausencia de actos humillantes se asocian con la masculinidad sana.

### 7.3.2. Importancia de variables

Para concluir con el modelo y proporcionar conclusiones más robustas, se procede a realizar un ranking de las variables más importantes del modelo.

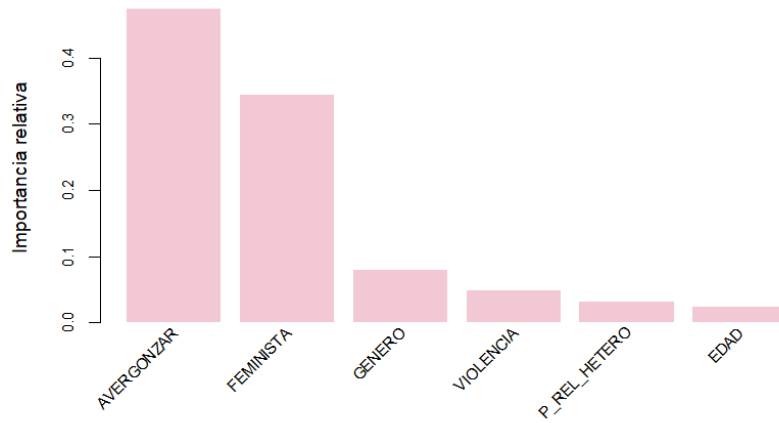


Figura 41: Importancia relativa variables árbol

En el gráfico anterior (Figura 40) se muestra la importancia relativa de las variables utilizadas en el árbol de decisión creado. Se observa que las variables más importantes son AVERGONZAR Y FEMINISTA. Esto nos dice, que la percepción de la masculinidad estará fuertemente definida por los comportamientos sociales e ideológicos del individuo. El resto de las variables incluidas en el modelo, siendo su importancia menor que las anteriores, son GENERO, VIOLENCIA, P\_REL\_HETERO y EDAD, en ese orden de importancia.

## 8. ¿Es la masculinidad frágil perjudicial para los hombres?

Para responder a esta pregunta, lo primero será definir qué es la masculinidad frágil. Este término hace referencia a la vulnerabilidad de la identidad masculina tradicional, cuando esta, se ve cuestionada o desafiada. Implica la creencia arraigada de que los hombres deben ajustarse a ciertos estereotipos de género, como la fortaleza emocional o la agresividad, percibiendo cualquier desviación de estas normas como una amenaza a su rol masculino (Soriano, 2023).

En este trabajo se entiende como masculinidad frágil aquella representada por la masculinidad tóxica (clúster 1), ya que representa la adhesión a roles de género tradicionales. Por lo tanto, volviendo a la pregunta, ¿perpetuar pensamientos tradicionales sobre la masculinidad, perjudica a los hombres?

Los datos evidencian que el grupo de masculinidad tóxica manifiesta mayores niveles de presión social para ajustarse a estereotipos tradicionales, como ejercer violencia para defender ideas, relacionarse únicamente con personas heterosexuales o consumir alcohol y drogas. Lo que implica una carga constante por cumplir con un ideal de masculinidad rígido. Además, los resultados del modelo de regresión muestran que la probabilidad de pertenecer a este grupo aumenta cuanto más frecuente es el uso de la violencia o la humillación hacia otras personas. Todo ello demuestra que perpetuar estos pensamientos, no solo impone sobre los hombres una presión que afecta a sus acciones y conductas, sino que afecta a las personas víctimas de esta violencia.

Por otra parte, los hombres con esta percepción tradicional presentan puntuaciones más bajas en indicadores de bienestar personal o percepción de ser uno mismo en el día a día. Por el contrario, el tener amistades con las que hablar de problemas personales aumenta significativamente la probabilidad de pertenecer al perfil de masculinidad sana, lo que indica un aislamiento emocional en el perfil tóxico. Asimismo, presentan una mayor frecuencia de problemas psicológicos y una mayor presión percibida para actuar y pensar como su grupo de amigos. Esto dificulta la libertad de expresión emocional y toma de decisiones autónomas. Esta situación, limita la capacidad de desarrollar relaciones basadas en la confianza y vulnerabilidad, elementos esenciales para el bienestar emocional.

A su vez, este grupo también se relaciona con sentir mayor presión por mostrar fortaleza en momentos de vulnerabilidad, reforzando que la masculinidad tradicional impide la expresión libre de emociones.

En conclusión, el perfil de masculinidad tóxica se asocia con mayor sufrimiento psicológico, mayor presión social y dificultad para establecer vínculos emocionales. Frente a ello, los modelos de masculinidad sanos se asocian con mayor bienestar personal y redes de apoyo afectivo.

Por lo tanto, perpetuar la masculinidad tradicional sí perjudica a los hombres, tanto en lo individual como en lo relacional.

## 9. ¿Hay diferencias entre hombres y mujeres en cuanto a la percepción de la masculinidad?

A lo largo del análisis, se han identificado diferencias significativas entre la percepción de la masculinidad entre hombres y mujeres.

En primer lugar, las mujeres muestran un mayor rechazo hacia los estereotipos tradicionales. Este patrón se observa en las distribuciones de respuesta, puesto que tienden a posicionarse en contra de las afirmaciones asociadas con la masculinidad hegemónica.

Analizando la distribución de los clústeres, se observa que aproximadamente el 70% de los individuos que pertenecen a los perfiles tradicional y neutro, son hombres. En el caso del perfil sano, este valor disminuye al 35%. Esta diferencia, evidencia que los hombres tienden a adoptar modelos más tradicionales.

Los resultados obtenidos en el modelo de regresión refuerzan este análisis, ya que el hecho de ser mujer aumenta significativamente la probabilidad de pertenecer al perfil de masculinidad sana.

Estas diferencias pueden explicarse por varios factores, desde edades tempranas, a los hombres se les enseña a reprimir emociones, competir y demostrar poder, mientras que a las mujeres se les exige el cuidado, empatía y expresión afectiva. Casi el 70% de los individuos, afirma haber crecido con un entorno familiar que definía estos comportamientos como los “correctos”. Esto contribuye al desarrollo emocional de las mujeres, y dificulta a los hombres a cuestionar los modelos tradicionales.

Además, los hombres experimentan beneficios derivados de la desigualdad de género, lo que puede llevar a que muestren mayor rechazo por la lucha del feminismo y la búsqueda de la igualdad, al percibir estos movimientos como una “amenaza” (Connell, 1995). Sin embargo, omiten que estos movimientos no buscan sustituir una dominación por otra, sino que benefician al conjunto de la sociedad.

No obstante, es importante señalar que una parte de los hombres se ubica en el perfil de masculinidad sana lo que indica que hay una evolución social en marcha.

En resumen, el análisis evidencia que las mujeres presentan una percepción más abierta e igualitaria respecto al concepto de masculinidad, mientras que los hombres, reproducen con más frecuencia creencias tradicionales. Esta diferencia tiene implicaciones con el bienestar emocional, tanto de los propios hombres como para quienes les rodean.

## 10. ¿Cómo perciben los jóvenes la masculinidad?

El estudio realizado, permite descubrir cómo perciben la masculinidad los jóvenes de España. En un contexto social, cada vez más marcado por el debate en torno a la igualdad de género, se muestra que la masculinidad es un concepto que se manifiesta de formas diversas y con distintos niveles de adhesión a estereotipos y comportamientos tradicionales.

Uno de los hallazgos más relevantes, es que aún persiste un grupo significativo de jóvenes que reproduce los pensamientos tradicionales, a pesar de haber una mayoría de individuos que cuestionan y rechazan la masculinidad hegemónica.

El análisis clúster permitió segmentar a los jóvenes en tres perfiles: masculinidad tóxica, neutra y sana. Estos resultados muestran tres formas de percibir y entender la masculinidad. Una perpetuando los pensamientos tradicionales y tóxicos, otro modelo crítico hacia la masculinidad hegemónica buscando perfiles más igualitarios y una tercera que comparte características con las dos anteriores.

El grupo mayoritario, casi el 58% de la muestra, pertenece al grupo de masculinidad sana, lo que supone una buena noticia ya que se observa una evolución social. Este perfil se caracteriza por una mayor libertad de expresión emocional, menor frecuencia de actos de riesgo, como la violencia o la humillación, así como una menor presión por realizar acciones como el consumo de drogas o perpetuar pensamientos heteronormativos. Asimismo, este grupo se asocia con el feminismo, ya que la probabilidad de pertenecer al perfil sano aumenta significativamente si el individuo pertenece a esta movimiento.

En contraste, el grupo asociado a la masculinidad tóxica mantiene esa visión tradicional de este concepto. Reproduce creencias sobre el constructo tradicional de la masculinidad, como asociar la violencia como la característica principal del hombre o la heterosexualidad obligatoria. Este grupo, además de mostrar mayor presión por cumplir estas “normas”, presentan indicadores más bajos de bienestar emocional. Este aspecto es clave, ya que este vínculo entre percepción de la masculinidad y bienestar demuestra que la masculinidad tradicional es limitante y dañina en términos de salud mental, aislamiento y presión social. Llama la atención, que en este grupo se encuentra la mayor proporción de personas pertenecientes a clases sociales altas. De esta manera se refleja que en los contextos socioculturales más privilegiados se perpetúan los roles tradicionales.

El perfil intermedio, denominado masculinidad neutra, refleja un rechazo parcial hacia algunos aspectos tradicionales, pero sigue perpetuando otros, lo que sugiere una evolución social en proceso.

En conjunto, este análisis evidencia que la juventud española se encuentra en un proceso de cambio social, reflejando una percepción sobre el concepto de la masculinidad mayoritariamente sano. Sin embargo, aún persisten jóvenes que mantienen esos perfiles tóxicos, especialmente entre los hombres. Esta evolución tiene un impacto sobre la calidad de vida, la libertad emocional y el desarrollo personal.

Por tanto, promover discursos y políticas que favorezcan una masculinidad más sana e impulsar una educación basada en el feminismo, es más que un objetivo de equidad social.

## Conclusión

Este estudio ha abordado la percepción de la masculinidad en la juventud española permitiendo identificar diferentes perfiles mediante técnicas como el análisis clúster, modelo de regresión logística multinomial y árboles de decisión. A lo largo del estudio se han identificado tres formas de percibir la masculinidad: tóxica, neutra y sana, y se ha analizado cómo la percepción de este concepto se relaciona con diversos factores como el género, la salud mental o la presión social.

Se han observado resultados que muestran una tendencia hacia el cambio, revelando que la mayoría de los jóvenes percibe la masculinidad de una manera más igualitaria. Sin embargo, aún se mantienen pensamientos y actitudes que ralentizan el cambio.

Por tanto, aún queda un largo camino por recorrer para conseguir un modelo de masculinidad plenamente igualitaria y libre de estereotipos. Como propone Octavio Salazar en *El hombre que no deberíamos ser*, no basta con cuestionar el concepto tradicional, hace falta deconstruir la masculinidad hegemónica y construir un nuevo modelo que sea beneficioso para el conjunto de la sociedad.

## Bibliografía

- Ajuntament de Barcelona. (s.f.). *Construcción social de las masculinidades*. Recursos Pedagògics.  
<https://ajuntament.barcelona.cat/recursospedagogics/es/masculinidades/construccion-social-de-masculinidades>
- Bussi, J., Hernández, L., Marí, G., Méndez, F., & Mitas, G. (2018). *Visualización y métodos de imputación de datos faltantes en la Encuesta de Gasto de los Hogares*. Vigésimoterceras Jornadas "Investigaciones en la Facultad" de Ciencias Económicas y Estadística. Instituto de Investigación Teórica y Aplicada, Escuela de Estadística. [Archivo PDF].
- Calviño, A. (2024). *Tema 2: Modelos de elección discreta y preferencia multidimensional*. Técnicas de segmentación y tratamiento de encuestas. [Archivo PDF]
- Calviño, A. (2024). *Tema 3: Árboles de clasificación y regresión y bosques aleatorios*. Técnicas de segmentación y tratamiento de encuestas. [Archivo PDF]
- Connell, R. W. (1995). *Masculinities*. University of California Press.
- De Wu, J. (2021, octubre 28). Damavis Blog - Data - Machine Learning – Visualization. *Análisis de Componentes Principales: Introducción matemática*.  
<https://blog.damavis.com/analisis-de-componentes-principales-una-breve-introduccion-matematica/>
- Panades, R. (2022, noviembre 26). *Masculinidad hegemónica: 15 ejemplos, definición y estudios de casos*. <https://ejemplode.online/masculinidad-hegemonica-15-ejemplos-definicion-estudios-de-casos/>
- Melian, V. (2025, marzo 1). *Construcción social de género: definición*.  
<https://factorhome.es/construccion/construccion/construccion-social-de-genero-definicion/>
- Instituto Nacional de Estadística (INE). (2023). *Tasa de mortalidad por lesiones debidas a accidentes de tráfico*. <https://ine.es/jaxi/Tabla.htm?tpx=46689&L=0>
- Salazar, O. (2018). *El hombre que no deberíamos ser: La revolución masculina que tantas mujeres llevan siglos esperando*. Planeta.
- Soriano, J. (2023, octubre 6). *Masculinidad Frágil: ¿los hombres deben deconstruirse?*. Portal Psicología y Mente. <https://psicologiymente.com/social/masculinidad-fragil>
- Patro, R. (2021, enero 25). *Cross-validation: K-fold vs. Monte Carlo*.  
<https://towardsdatascience.com/cross-validation-k-fold-vs-monte-carlo-e54df2fc179b/>
- Wolf Diseño Web. (2025, abril 19). *Algoritmo de k-means: qué es y cómo funciona*.  
<https://wolfdisenoweb.com/2025/04/19/algoritmo-de-k-means/>