

# Estudio de los resultados del Campeonato Europeo de Halterofilia

Cliente: *Lift & Lead, Crossfit centers*

*Overview* del análisis

Alicia Penalba

Mayo de 2024

# CONTENIDO\*

## 01 Contexto

- Campeonato de Halterofilia
- CrossFit
- Definición del problema

## 02 Limpieza y tratamiento de los datos

## 03 Análisis: 2019 y 2020

- Exploración de datos
- Preguntas a resolver

## 04 Análisis ampliado : 2019 a 2024

- Preguntas a resolver
- Medallas españolas

## 05 Resultados y recomendaciones



# 01 Contexto: Campeonato de Halterofilia

## Datos de interés

- Se empezó a celebrar en **1896** y sigue las reglas y regulaciones establecidas por la **Federación Internacional de Halterofilia (IWF)**.
- Los competidores se dividen en varias **categorías de peso**, tanto en género **masculino** como **femenino**.
- Se otorgan medallas de **oro, plata y bronce** en cada categoría de peso y por cada género.
- Los atletas compiten en **dos ejercicios** distintos, con 3 intentos en cada uno de ellos.



# 01 Contexto: Campeonato de Halterofilia

## Datos de interés

- **Arranque o arrancada:** los atletas levantan la barra desde el suelo hasta por encima de la cabeza en un único movimiento continuo.
- **Envi3n o Dos tiempos:** los atletas levantan la barra desde el suelo hasta los hombros y luego por encima de la cabeza en dos movimientos separados.
- La puntuaci3n es el **total de kg levantados** en cada uno de los ejercicios.

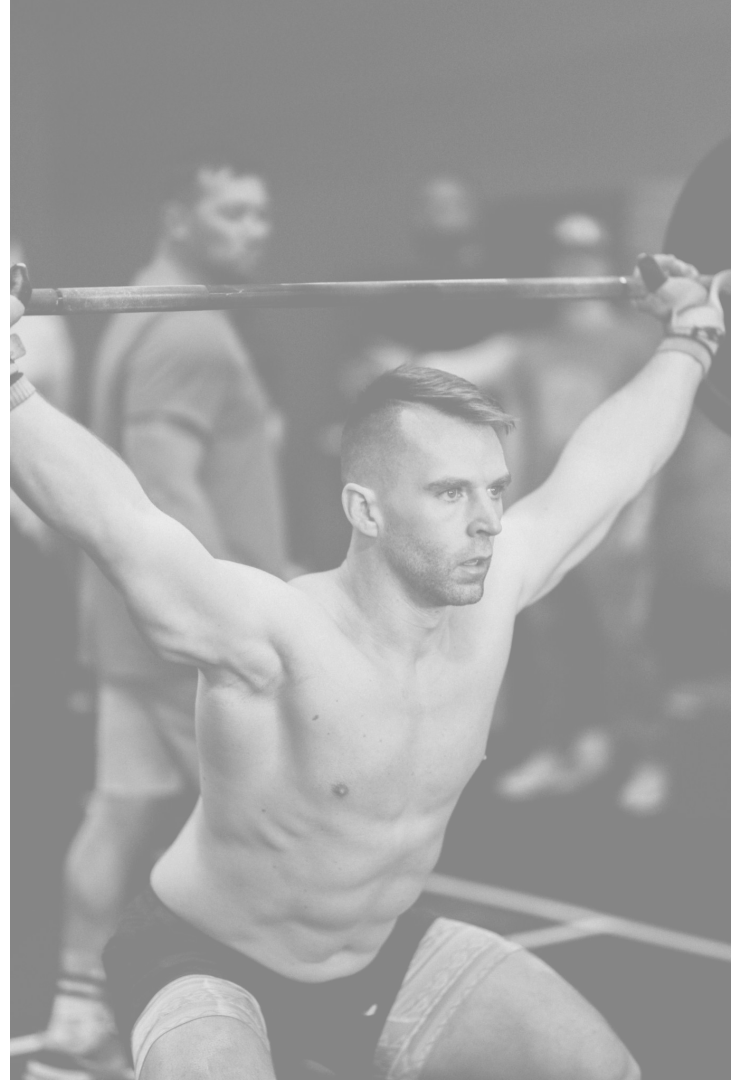




# 01 Contexto: Crossfit

## Datos de interés

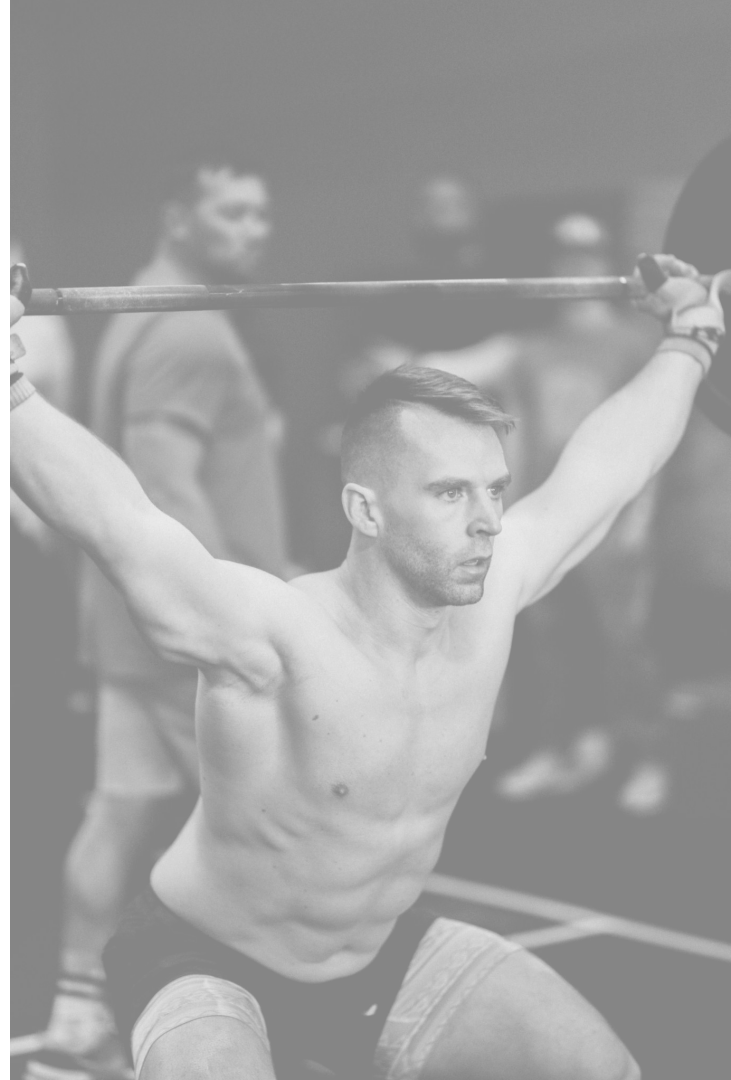
- **CrossFit** es un entrenamiento registrado, propiedad de **CrossFit Inc.** que conecta movimientos de diferentes disciplinas, tales como la **halterofilia** o los **entrenamientos gimnásticos**.
- Su creador es **Gress Glassman**, estadounidense que durante su trayectoria deportiva como gimnasta, descubrió el uso de mancuernas y barras de levantamiento para realizar ejercicios de fuerza y decidió desarrollar entrenamientos de creación propia de alta intensidad.
- La compañía se fundó en 1996, y para **2018** ya había unos **15.500 gimnasios de CrossFit en 162 países**.



# 01 Contexto: Crossfit

## Datos de interés

- **España** es uno de los **países europeos** con más centros de Crossfit, junto con Inglaterra e Italia.
- Los entrenamientos CrossFit o **“WOD”** (work of the day) duran aproximadamente 60 minutos e incluyen:
  1. **Calentamiento**
  2. **Técnica/Fuerza “A”**: se practica la técnica de diferentes ejercicios en los que se desarrolla la fuerza, la técnica y la potencia con más peso.
  3. **Parte “B”**: Es el momento más intenso del entrenamiento en el cual se baja la cantidad de peso en comparación con la parte “A” y se busca aumentar la capacidad respiratoria.



# 01 Contexto: definición del problema

El cliente plantea tres preguntas **concretas** respecto a los datos del Campeonato de Halterofilia. Tienen un **alcance definido**, creen que tienen **valor** para su compañía y en base a eso podrán definir mejor una estrategia para tomar **acción** en la apertura de centros de Crossfit:

¿**Cuántas medallas** de cada tipo ha ganado cada país?

¿Qué **país** ha logrado **mayor equidad** en términos de **éxito** entre atletas masculinos y femeninos?

¿En qué **país** es menor la diferencia entre el **promedio de puntos totales** por género?

# 01 Contexto: definición del problema



- Analizamos los **datos de 2019 y 2020**. Posteriormente realizamos *web scraping* para obtener los de **2022, 2023 y 2024**. No se incluyen datos de 2021, ya que por **COVID el campeonato de 2020 se anuló**. Por lo tanto, los datos de 2020 se generaron en 2021, pero corresponden al campeonato atrasado por la pandemia.
- Nuestro análisis inicial se centra en una muestra de **120 resultados**, con **21 países** representados.
- El segundo análisis ampliado recoge un total de **300 resultados**, con **26 países** representados.
- Variables numéricas: **Arrancada, Dos tiempos, Total**.
- Variables categóricas: **Género, Categoría, Medalla, Nombre, Apellido, País**



## 02 Limpieza y tratamiento de los datos

### Transformación

Eventos	Oro	Plata	Bronce	Año	Género*
55 kg (04.04)	Anguel Rusev Bulgaria 111 + 147 = 258[n 1]	Valentin Iancu Rumania 108 + 140 = 248[n 1]	Dmytro Voronovsky Ucrania 110 + 137 = 247[n 1]	2020	Masculino



Género*	Categoría	Fecha	Medalla	Nombre	Apellido	País	Arrancada	Dos tiempos	Total
Masculino	55	2019	Oro	Anguel	Rusev	Bulgaria	111	147	258

### Técnicas usadas:

- **Cambio de orientación** de la tabla: creación de la columna categórica 'Medalla'
- **Expresiones regulares** para la identificación de patrones
- **Separación de elementos** de una columna en varias: 'Nombre', 'Apellido', 'País'...

\* Disclaimer:

Los datos se basan en tablas de resultados de la competencia oficial, que sólo incluyen masculino y femenino sin considerar otros géneros.

## 02 Limpieza y tratamiento de los datos

### Modificaciones específicas e información relevante

- **Rusia y Bielorusia** se excluyeron de la competición en **2022 y 2023** por la guerra contra Ucrania. En **2024** se deja participar a atletas de Bielorusia bajo la nomenclatura **AIN** (atletas internacionales neutros). Modificamos AIN por Bielorusia.
- 'Reino Unido' se ha renombrado como 'UK'.
- En los atletas con **nombres compuestos** o **dos apellidos**, unimos mediante guión bajo las dos primeras entradas de la columna 'Atleta'.



# 03 Análisis 2019 y 2020: Exploración de datos

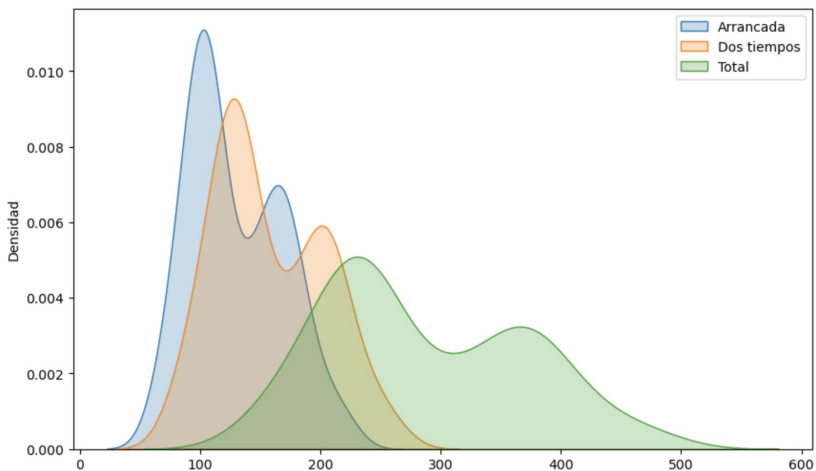
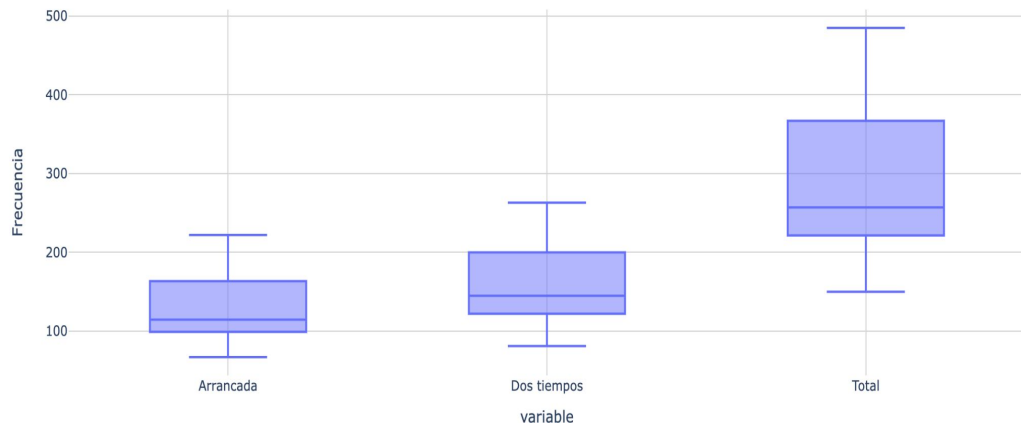
## Variables numéricas

### Arrancada, Dos tiempos, Total:

Representamos la distribución de las puntuaciones de dos maneras: **boxplot** y **gráfico de densidad**.

Podemos ver que los datos son correctos. Las **puntuaciones** corresponden a los **kg levantados** por los atletas en cada fase, por lo que los valores en el ejercicio de arrancada son siempre más bajos que en los de dos tiempos, debido a la naturaleza del ejercicio. El total es la suma de ambos.

En la arrancada, los atletas deben hacer el movimiento de una vez, por lo que el peso que pueden levantar es menor que en la segunda prueba.



## 03 Análisis 2019 y 2020: Exploración de datos

### Variables numéricas

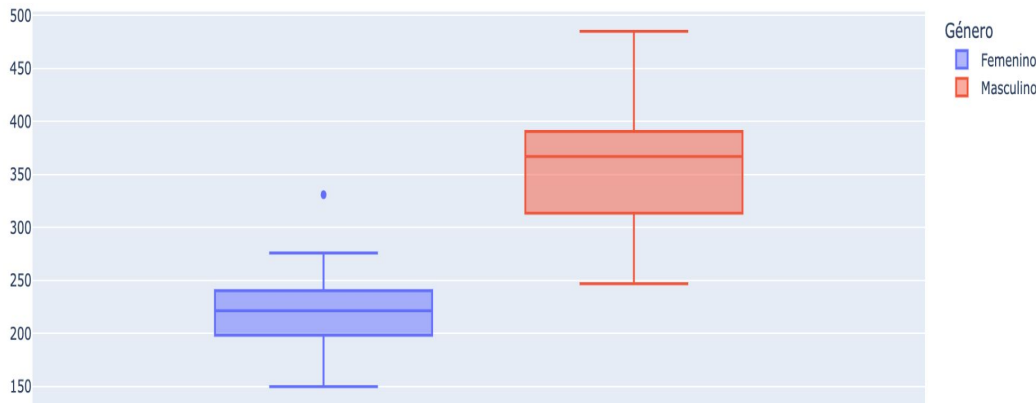
#### Total por género:

Las puntuaciones totales por género muestran que los atletas masculinos obtienen puntuaciones mayores. Esto es debido a que en términos generales pueden llegar a levantar más peso, que es lo que genera el valor de la puntuación.

De todos modos, observamos cómo hay atletas de género femenino que consiguen puntuaciones más altas que algunos atletas de género masculino. Hay un valor atípico, correspondiente a las puntuaciones totales de la rusa **Tatiana Kashirina en 2019**, que **diste tan sólo 36 kg de la puntuación mediana masculina**.

Exceptuando este caso en concreto, no se observan valores atípicos en la muestra.

Distribución puntuaciones Totales por género

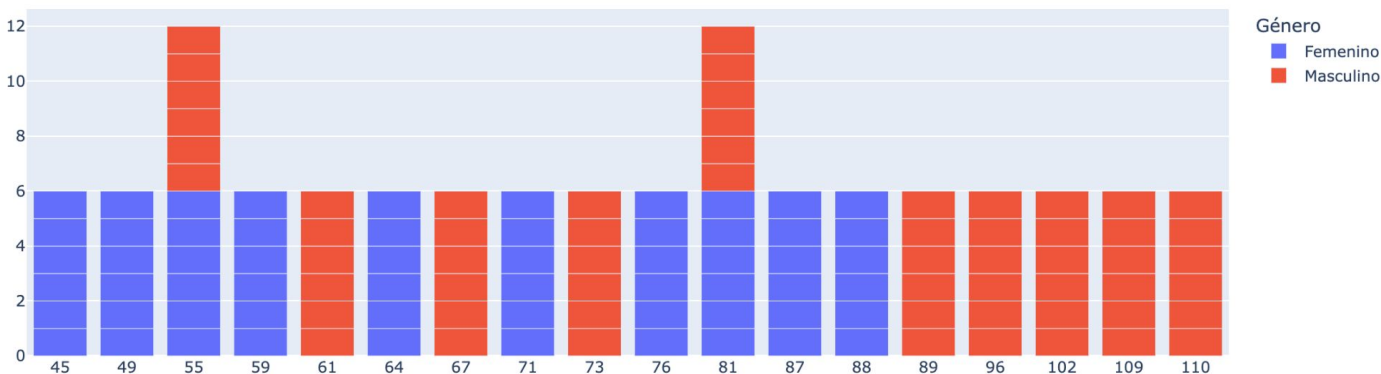


## 03 Análisis 2019 y 2020: Exploración de datos

### Variables categóricas

#### Categoría de peso y género

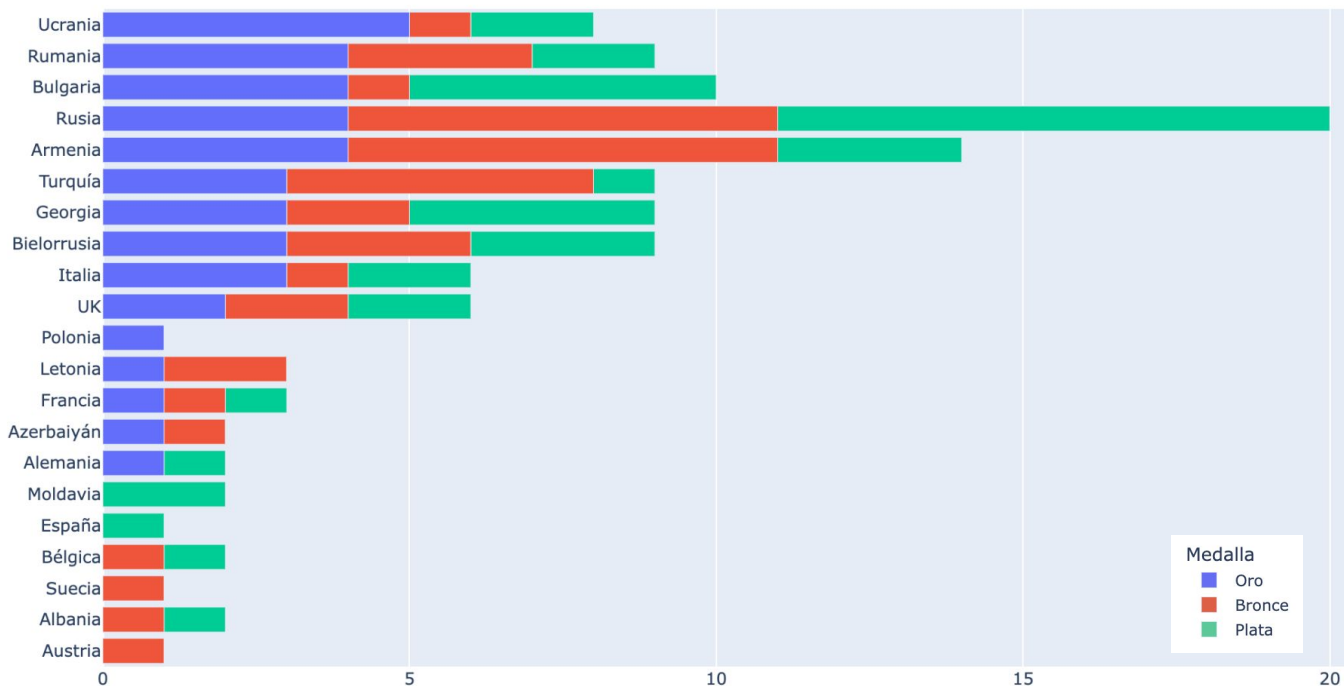
- Hay un **total de 18 categorías**. Las que son de **88 kg y 110 kg** son realmente +87 y +109, pero para facilitar el análisis hemos quitado el + y añadido 1 kg para incluir a los atletas que se encuentran en los límites superiores de peso.
- La distribución de pesos por categoría es menor en el caso femenino, es decir, hay menos diferencia de peso entre ellas. Normalmente, la distancia entre **categorías femeninas va de 4 a 6 kg**.
- En el caso de las **categorías masculinas, se distancian entre los 6 y 8 kg**.
- Las categorías en las que coinciden ambos géneros son **55 kg y 81 kg**.





### 03 Análisis 2019 y 2020: Preguntas a resolver

¿Cuántas medallas de cada tipo ha ganado cada país?



Los países con un **mayor número** de medallas son **Rusia y Armenia**, con un total de **20 y 14**, respectivamente. Empatan en oros y bronce, pero en plata Rusia consigue llevarse 9, frente a las 3 de Armenia.

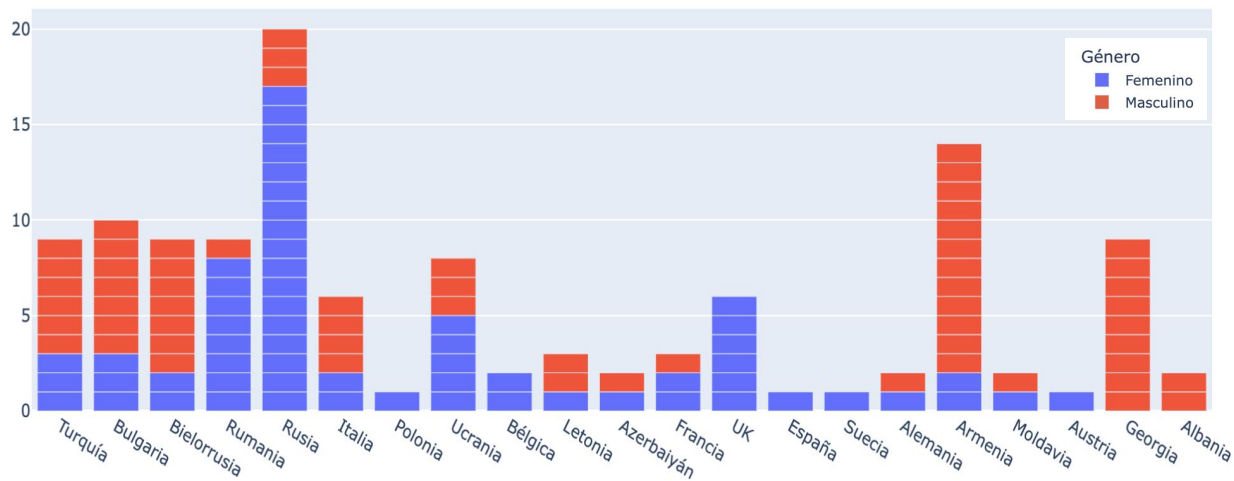
Les siguen otros países de la zona oriental y del este, como **Bulgaria** con 10 medallas y **Turquía, Georgia y Bielorrusia** con un total de 9. **Ucrania** obtiene la 8a posición con 8 medallas.

Los países con **menos victorias** son **Polonia, España, Suecia y Austria**, con 1.

Por lo tanto, se evidencia la **importancia** de la halterofilia en la **zona este y sudeste de Europa**.

### 03 Análisis 2019 y 2020: Preguntas a resolver

¿Qué país ha logrado mayor equidad en términos de éxito entre atletas masculinos y femeninos?



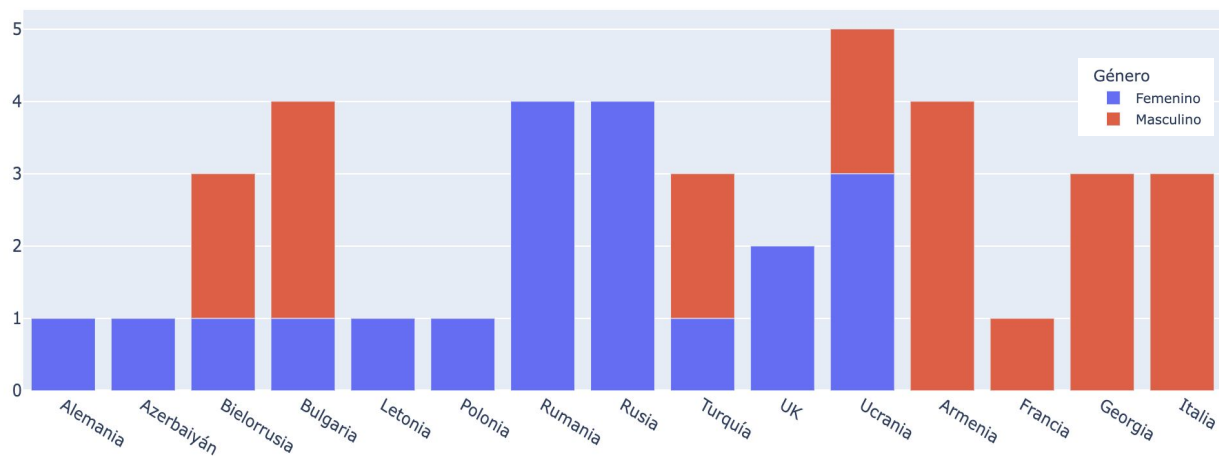
Consideramos como **éxito la obtención de medallas** y como **equidad la menor diferencia** entre el total de medallas por género.

Para responder a esta pregunta hemos realizado una **visualización del número de medallas por género** y también una **tabla de contingencia** para calcular el valor de la equidad en cada uno de los países. Como más cercano a 0 ese valor, mayor equidad.

Observamos cómo en este caso, **los países con mayor equidad son Azerbaiyán, Alemania y Moldavia**, ya que cuentan con las mismas medallas en masculino y femenino.

### 03 Análisis 2019 y 2020: Preguntas a resolver

¿Qué país ha logrado mayor equidad en términos de éxito entre atletas masculinos y femeninos (medallas de oro)?



Si visualizamos los mismos datos pero sólo con las **medallas de oro**, vemos que no hay ningún país que obtenga la misma proporción de medallas tanto en masculino como en femenino.

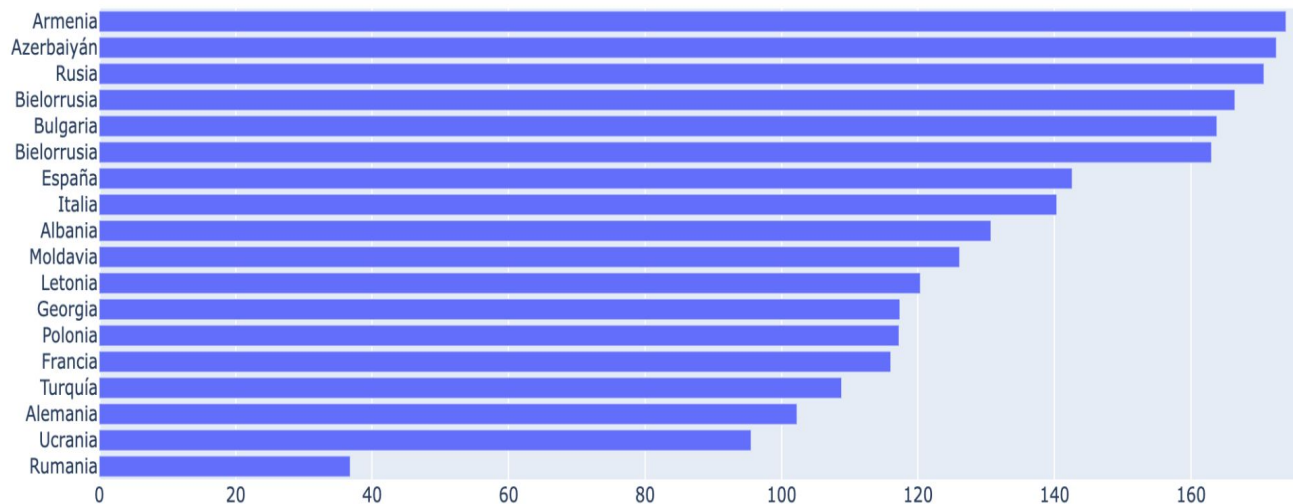
De hecho, tan sólo **4 países (Bielorrusia, Bulgaria, Turquía y Ucrania)** consiguen **medallistas de oro** en **ambos géneros**.

En **género masculino** destacan **Armenia, Georgia e Italia**.

En **género femenino**, los países más reconocidos son **Rusia y Rumanía**, con 4 medallas.

### 03 Análisis 2019 y 2020: Preguntas a resolver

¿En qué país es menor la diferencia entre el promedio de puntos totales por género?



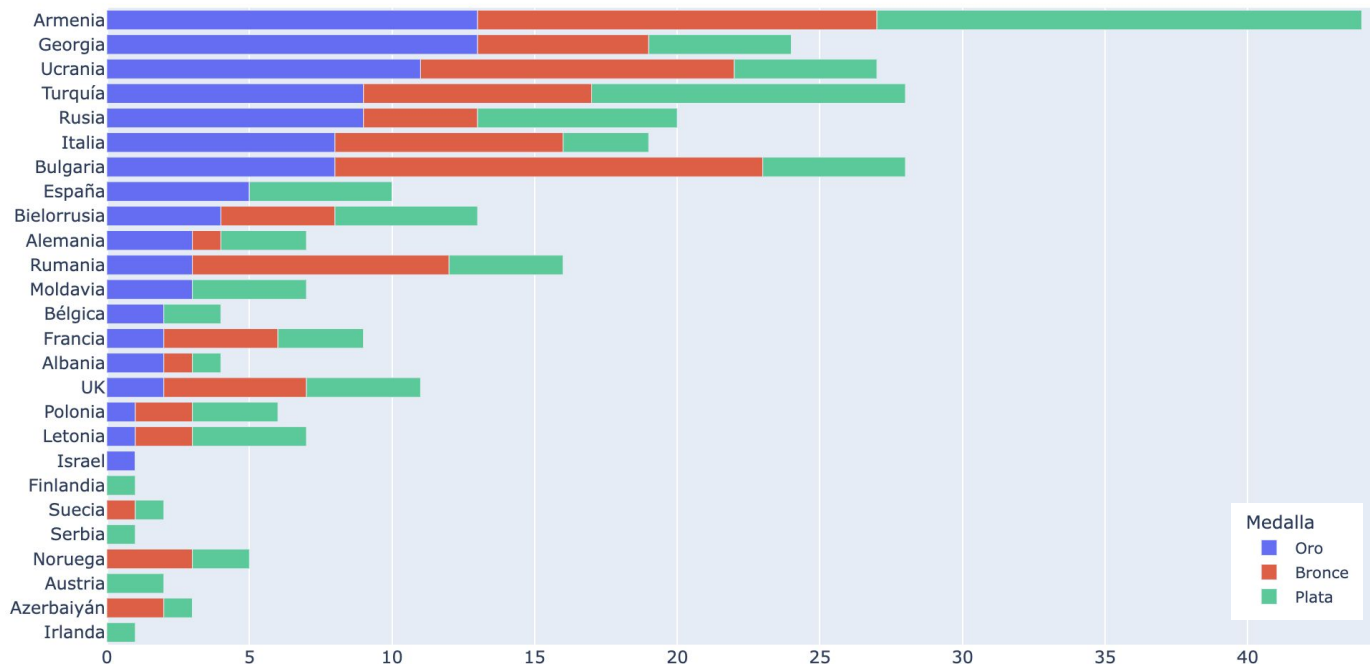
Hemos calculado el promedio por cada género de las puntuaciones totales y hemos restado ambos valores. Los países con representación en sólo uno de los géneros no aparecen en el gráfico.

Como mayor es el valor de la diferencia, esto nos indica que las puntuaciones entre atletas masculinos y femeninos son más dispares.

El país con una diferencia menor es Rumanía, con tan sólo 35 puntos entre promedios. Esto puede indicar que sus atletas femeninos y masculinos compiten en categorías de peso similar.

## 04 Análisis ampliado 2019-2024\*: Preguntas a resolver

¿Cuántas medallas de cada tipo ha ganado cada país?



**Armenia** consigue en este período posicionarse como el **primer país** en número de medallas, mientras que **Rusia está en 6a posición** con las mismas medallas que en el análisis de 2020 a causa de su prohibición de participación por la guerra.

**Bulgaria, Ucrania, Turquía y Georgia** siguen entre las primeras posiciones.

**Italia**, que en el análisis anterior estaba a la par que Reino Unido ha incrementado su número de medallas considerablemente, distanciándose de los británicos.

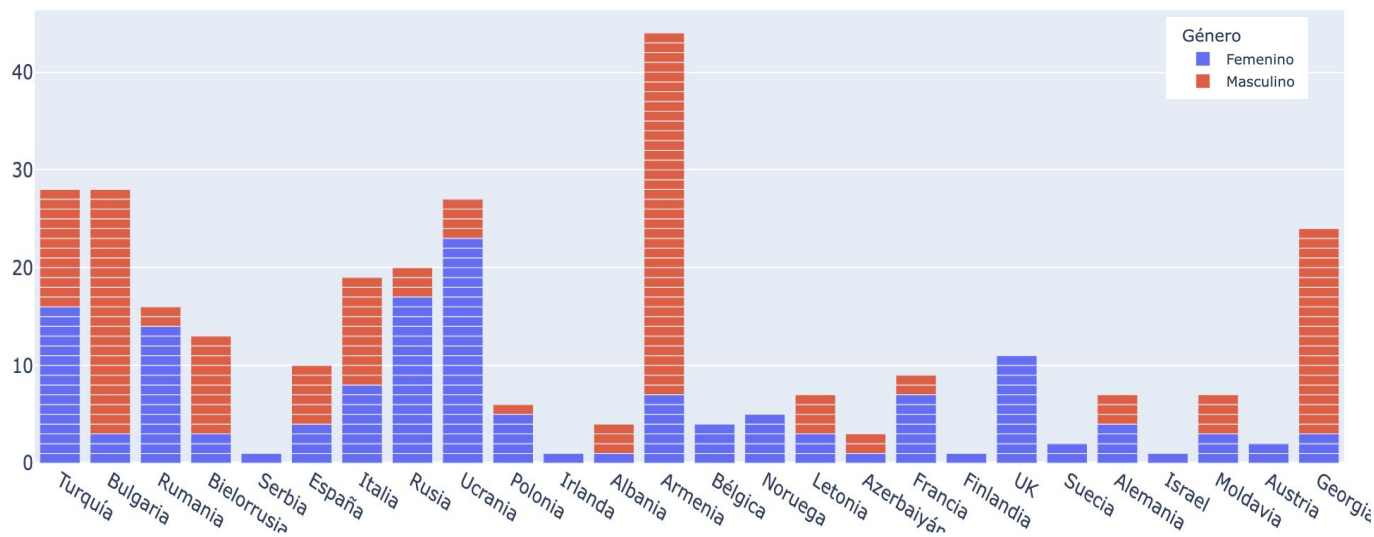
**España** también **despunta** con 10 medallas y consigue escapar de esas últimas posiciones, que ahora son de Finlandia, Irlanda, Serbia e Israel.

\*El análisis exploratorio es idéntico al de la primera parte, sólo que con mayor cantidad de datos. Al estar ya familiarizados con estos, pasamos directamente a las preguntas a resolver



## 04 Análisis ampliado 2019-2024: Preguntas a resolver

¿Qué país ha logrado mayor equidad en términos de éxito entre atletas masculinos y femeninos?



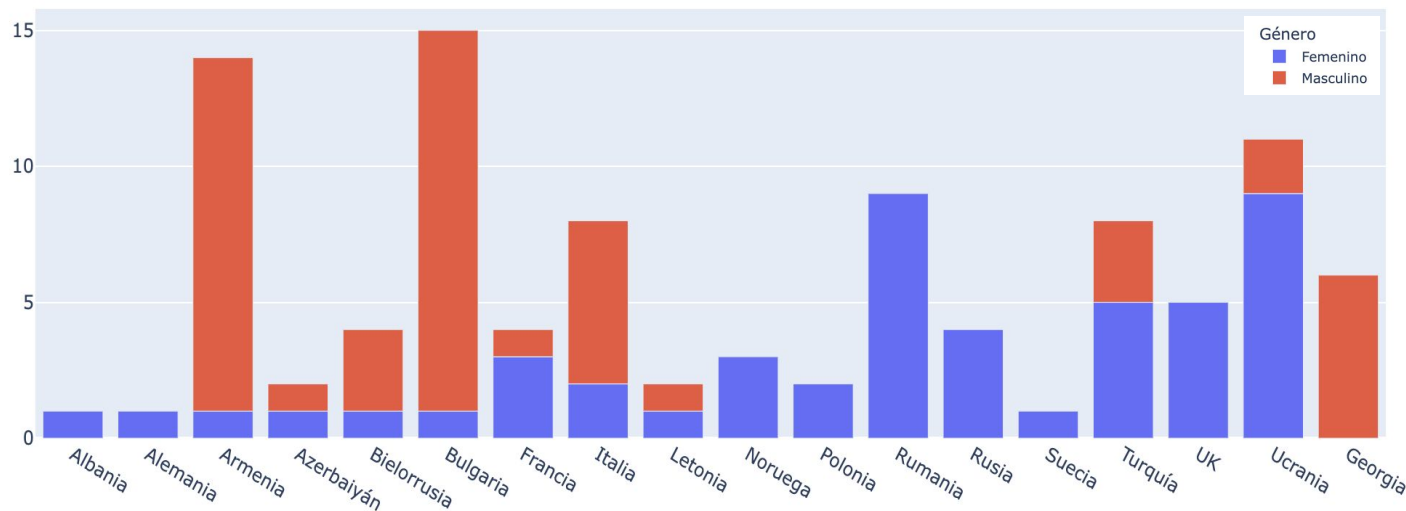
Realizamos el mismo proceso que en el análisis previo.

Observamos cómo en este caso, **los países con mayor equidad son Alemania, Moldavia, Letonia y Turquía**. Los tres primeros tienen total de 7 medallas, con una proporción 4-3 entre géneros. En el caso de Turquía, tiene 12 en masculino y 16 en femenino.

Los países con menor equidad son aquellos en los que sólo hay 1 género representado, como Serbia, Irlanda, Bélgica o Reino Unido..

## 04 Análisis ampliado 2019-2024: Preguntas a resolver

¿Qué país ha logrado mayor equidad en términos de éxito entre atletas masculinos y femeninos (medallas de oro)?

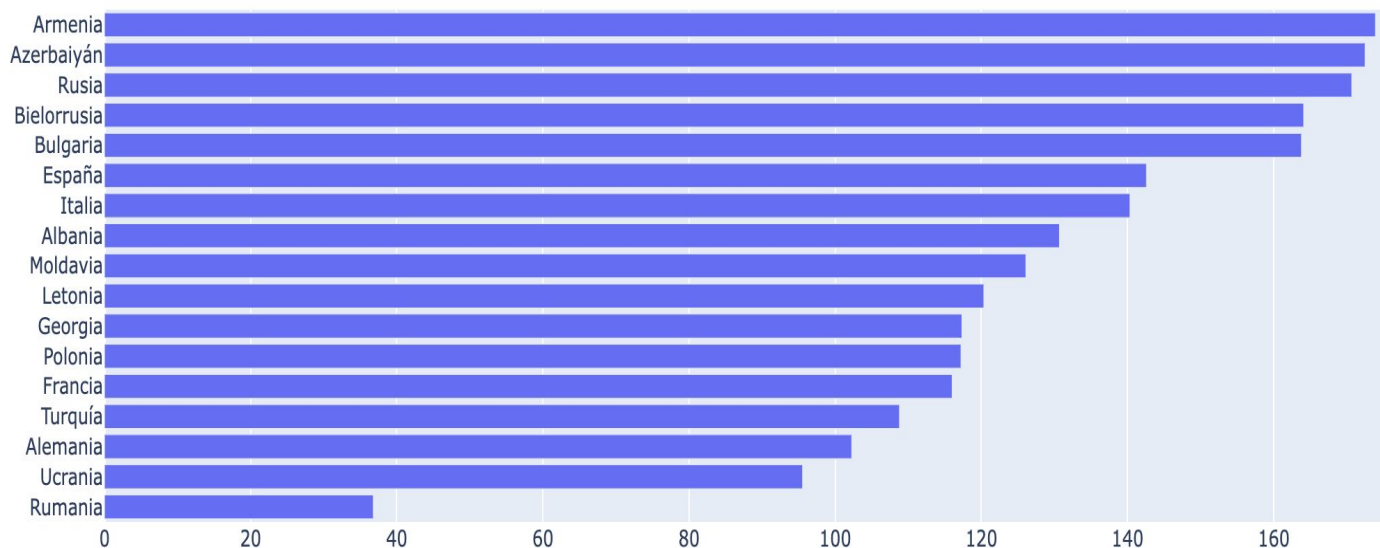


Si visualizamos los mismos datos pero sólo con las **medallas de oro**, en este caso **Azerbaiyán y Letonia** consiguen la mayor equidad, con **1 medalla de oro por cada género**.

Con la ampliación de datos, más países han conseguido oros en ambos géneros: Armenia, Azerbaiyán, Bielorrusia, Bulgaria, Francia, Italia, Letonia, Turquía y Ucrania.

## 04 Análisis ampliado 2019-2024: Preguntas a resolver

¿En qué país es menor la diferencia entre el promedio de puntos totales por género?



No se observan cambios muy significativos respecto al análisis anterior.

El país con una diferencia menor sigue siendo Rumanía, lo que puede indicar que sus atletas femeninos y masculinos compiten en categorías de peso similar.

# 04 Análisis ampliado 2019-2024: Medallas españolas

Esta tabla muestra un resumen de la información del medallero Español.

Las medallas se distribuyen del siguiente modo: **3 bronce y 1 plata en género femenino y 4 platas y 2 bronce en género masculino.**

**Marta García y Acorán Hernández** consiguieron repetir medalla en sus respectivas categorías durante dos años consecutivos.

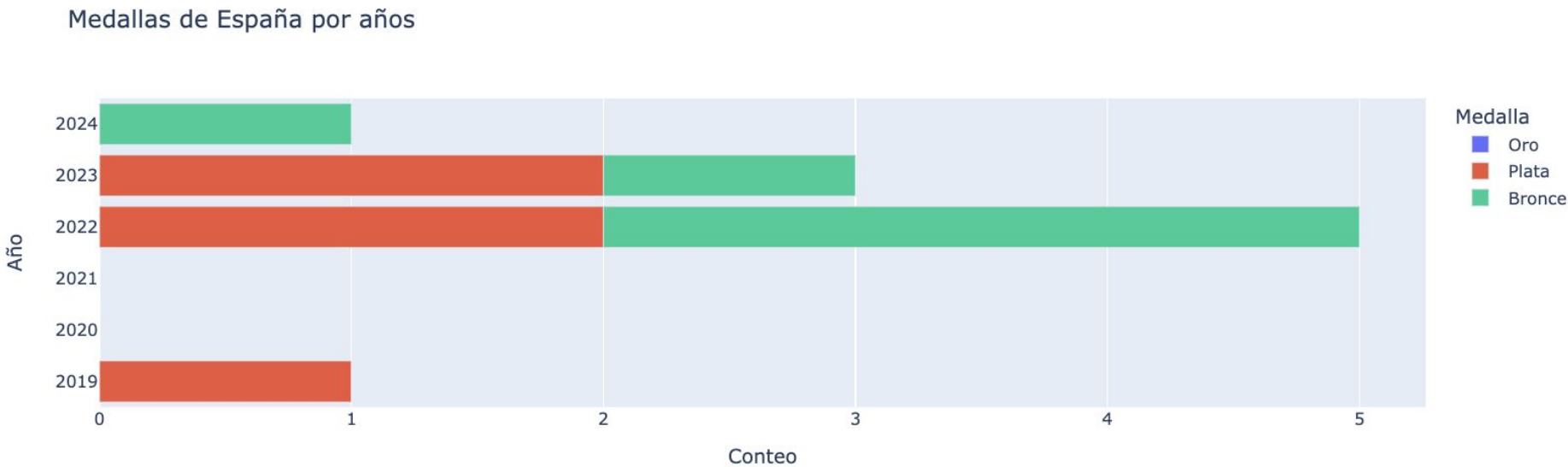
En **2019** se obtuvo una única medalla gracias a la atleta olímpica **Lidia Valentín**.

El incremento de medallas para España puede verse en la siguiente diapositiva.

Género	Categoría	Año	Medalla	Nombre y apellido
Femenino	45	2023	Bronce	Marta García
Femenino	45	2024	Bronce	Marta García
Femenino	49	2022	Bronce	María Giménez-Guervos
Femenino	76	2019	Plata	Lidia Valentín
Masculino	55	2022	Plata	Josué Brachi
Masculino	67	2023	Plata	Acorán Hernández
Masculino	67	2022	Bronce	Acorán Hernández
Masculino	73	2023	Plata	David Sánchez
Masculino	81	2022	Plata	Andrés Mata
Masculino	102	2022	Bronce	Marcos Ruiz

# 04 Análisis ampliado 2019-2024: Medallas españolas

El medallero español ha pasado de **1 victoria en 2019** a **10** en tan sólo **4 competiciones**. Necesitaríamos más datos para determinar cuáles han sido los posibles motivos de este incremento, pero el hecho de que Rusia y Bielorusia fueran excluidas de la competición parece haber tenido efectos positivos para España. Recordemos que son países referentes en este deporte, y que en 2019 y 2020 ocupaban las primeras posiciones.





# 05 Resultados y recomendaciones

## Resultados

- Los **países del este y sureste de Europa** han predominado claramente durante todo el período de análisis en la obtención de medallas.
- Parece ser que la **exclusión de Rusia y Bielorusia** de la competición en los años 2022 y 2023 ha permitido a otros países ampliar su medallero. Podrían ser los casos de España e Italia.
- **Alemania, Moldavia, Letonia y Turquía** destacan por ser los **países con una mayor equidad entre género masculino y femenino**. Es decir, la proporción de medallas es similar para ambos géneros.
- **Rumanía** destaca por ser el país con una menor diferencia de puntuación promedio entre géneros.

## 05 Resultados y recomendaciones

### Recomendaciones

1. Este análisis preliminar ha permitido sacar conclusiones interesantes sobre el peso de la halterofilia a nivel profesional en Europa, pero consideramos que sería de interés para nuestro cliente ampliar los datos para investigar si realmente hay una **correlación entre el auge del CrossFit y esta disciplina**.
2. Por ejemplo, podría ser de relevancia **analizar el peso del CrossFit por países**, y compararlo con sus resultados. Ver si en aquellos donde se obtienen más medallas también hay más adeptos a este deporte.
3. Como nuestro cliente está interesado en el **mercado español de centros de CrossFit**, le recomendamos hacer un estudio más exhaustivo de este deporte en el territorio.
4. Si quisiera seguir centrándose en la relación entre la halterofilia y el CrossFit, consideramos que una opción interesante sería contactar con los **medallistas españoles** para saber cuáles son sus **rutinas de entrenamiento** y saber si sería posible trabajar con ellos para realizar campañas publicitarias que pudieran **relacionar a la start-up** con los y las atletas.