

# 120 AÑOS DE HISTORIA DE LOS JUEGOS OLÍMPICOS

Garnica Cárdenas M.F.  
Morales García I.  
García Partida Y  
Marentes Camarillo C.  
Dayanne Alejandra Vidal

## INTRODUCCIÓN

Los juegos olímpicos son el máximo evento deportivo internacional, están inspirados en los juegos olímpicos de la antigüedad organizados en la antigua Grecia, la 1era edición de estos se llevo acabo en Atenas, Grecia.

Con los datos obtenidos de la base pretendemos dar a conocer cuáles han sido los deportes con más medallas, y cual es el competidor promedio ideal para ganar oro en cada deporte.

Tenemos una ligera hipótesis de que el factor edad influye en los ganadores de las preseas exceptuando obviamente a los atletas superlativos.

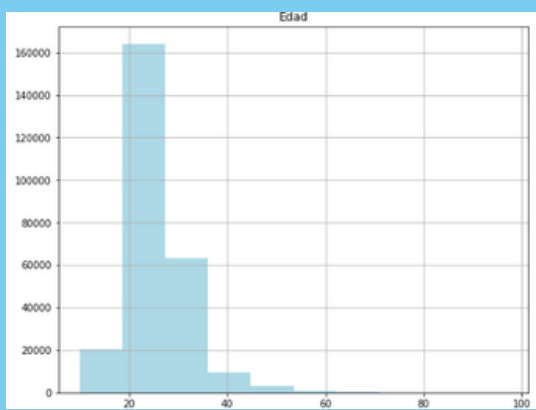
## OBJETIVO

### Principal:

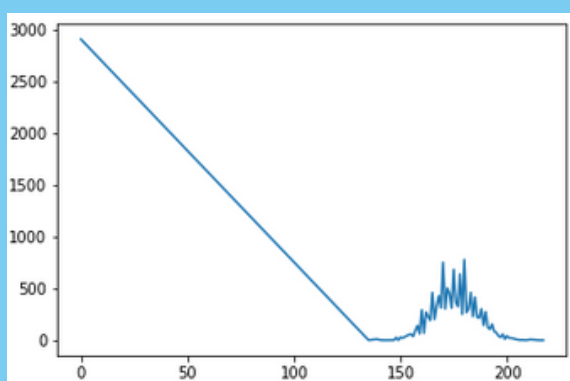
Dar a conocer las características específicas que han ayudado a los competidores a lograr colgarse de una presea.

### Secundario:

Conocer un promedio de edad de los competidores que logran colgarse una presea.



Se observa que la edad promedio es de 25 años.



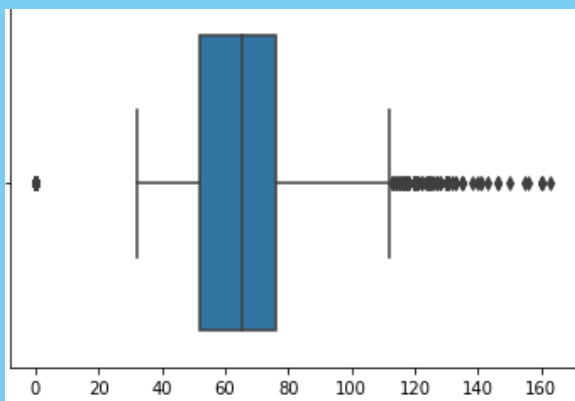
Se observa que la estatura promedio es 1.75m

## METODOLOGÍA

Se crearon dos bases de datos para prevenir la perdida de datos masiva. Una para el peso y la estatura y otra con los demás datos quitando peso y estatura.

Se eliminaron dos columnas "NOC" y "Games" ya que solo eran datos repetidos.

Ya al tener las 2 bases se eliminaron los 0's en edad, estatura y peso.



Se observa que el peso ideal es de 70.7kg

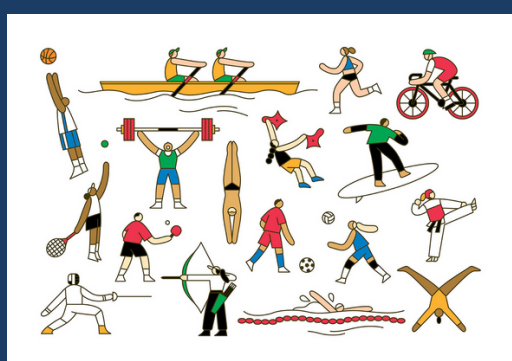
## RESULTADOS

En el modelo de bosques de arboles de decisión se procede a dividir los dato en un 80% para entrenamiento y el resto para validación. Así mismo se clasifican 6 condiciones estrechamente relacionadas a la caracterización de cada atleta como variables explicativas y se titula a la variable "Medal" como el valor a predecir. Los hiperparámetros del modelo se establecen en 50 arboles, 20 subdivisiones y 10 cortes por cada árbol con esto decimos que los resultados del modelo no son conclusivos, ya que si bien obtenemos un conjunto de métricas positivas, entre ellas un 77.46% de poder explicativo promedio en los distintos niveles contemplando las 6 categorías. Uno de los resultados más relevantes descansa en la jerarquización por relevancia de las variables explicativas, en donde destacan el equipo, el deporte y la edad; seguidas por el peso, la estatura y el sexo respectivamente. Lo anterior apunta a que los equipos o países se pueden apreciar como una entidad mas o menos homogénea en el desarrollo de sus atletas internos.

## CONCLUSIÓN

La edad, el peso y la estatura son descriptores fisionómicos de los atletas que podríamos presuponer que están correlacionados con su país de origen dado a la evidente diferenciación de las medias poblacionales de estas variables entre países. .

En conclusión, el modelo de bosques de arboles de decisión nos entrega un conjunto de interpretaciones relevantes para nuestro estudio. Finalmente, el modelo puede mejorarse a partir de un tratamiento distinto de sus variables, distintas simulaciones con configuraciones de sus hiperparámetros distintas, que quedan fuera de este ejercicio dada la carga de computo, así como un análisis de componentes principales para contrastar las hipótesis aquí propuestas.



## RECURSOS

