



Proyecto EDA

ANÁLISIS DE RENDIMIENTO EN JUGADORES DE LA RFEF TEMPORADA 23/24.

Nicolas Matías Muñiz Escalada – Bootcamp Data Science

Por que este tema?



Desde la película Moneyball me interesan las estadísticas y como se pueden aplicar el Big Data y análisis de datos en el mundo del deporte, a fin de dar ayuda a los distintos clubes/jugadores/entrenadores que estén interesados en aplicarlo

Hoy en día se esta utilizando a gran escala justo en el deporte que mas me apasiona, como es el Futbol.

Por eso, he escogido un Data Set que presenta estadísticas básicas de jugadores pertenecientes a la RFEF y comenzar a analizar aspectos básicos de rendimiento orientándome a los pensamientos que tengo con respecto al funcionamiento de este juego.

En este primer proyecto, introduzco mis nuevos conocimientos de como gestionar los datos, y lo enfrento con mucho trabajo.

Contexto



El data set contiene información detallada sobre jugadores de fútbol de la temporada 23/24 de la Real Federación Española de Fútbol (RFEF). Los datos incluyen atributos como:

- Nombre del jugador.
- Posición.
- Goles anotados.
- Asistencias.
- Partidos jugados y titularidades.
- Edad, altura, valor en el mercado y rating.

El objetivo es analizar el rendimiento de los jugadores para identificar patrones, destacar a los mejores de la temporada, y explorar diferencias significativas entre grupos definidos por posición y características físicas.

Planteo de Hipotesis



- **Posición por altura y edad:** Existen diferencias entre las estaturas de los jugadores según la posición que jueguen. Entiendo que para ser portero, necesitas ser mas alto, lo mismo para defensores, y delanteros. La edad también influye ya que en posiciones laterales o extremas, se necesita mas gente joven.
- **Conducta por posición:** Los jugadores que se encuentren ubicados detrás de la mitad de cancha excluyendo al portero (Defensores y Mediocampistas) son los que mas amonestaciones llevan en la temporada.
- **Goles en los delanteros centrales:** Si los delanteros hacen mas goles entonces recibirán mas raiting, y a su vez también jugaran mas partidos de titular.
- **Experiencia dentro del campo:** En general los equipos optan por tener jugadores con experiencia dentro del campo de juego, acompañando a los Jóvenes y Novatos.

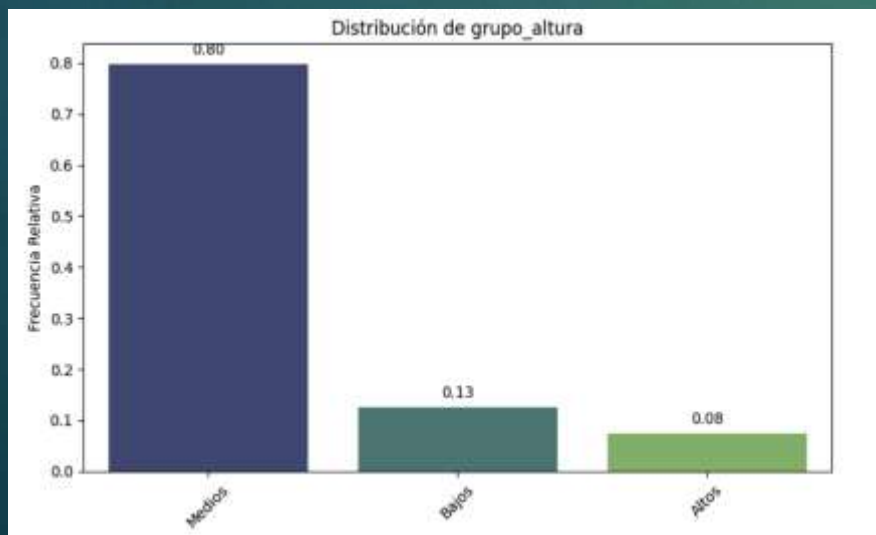
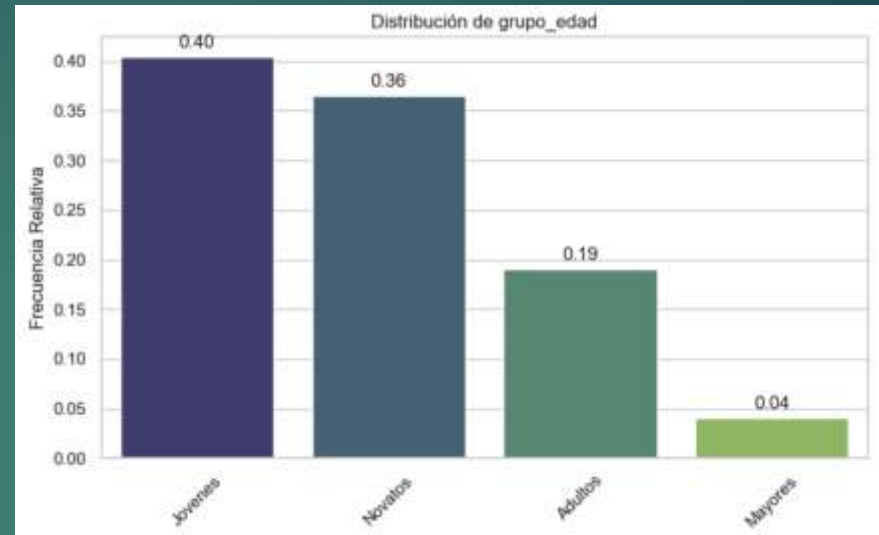
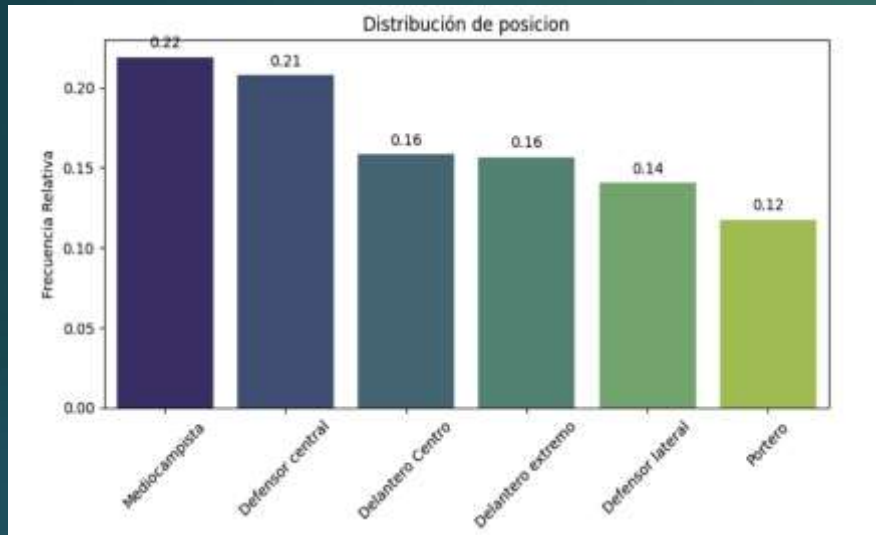
Analisis Univariante:



He presentado todas las variables clasificándolas cada una de ellas como categóricas o numéricas para luego comenzar con el análisis bivalente.

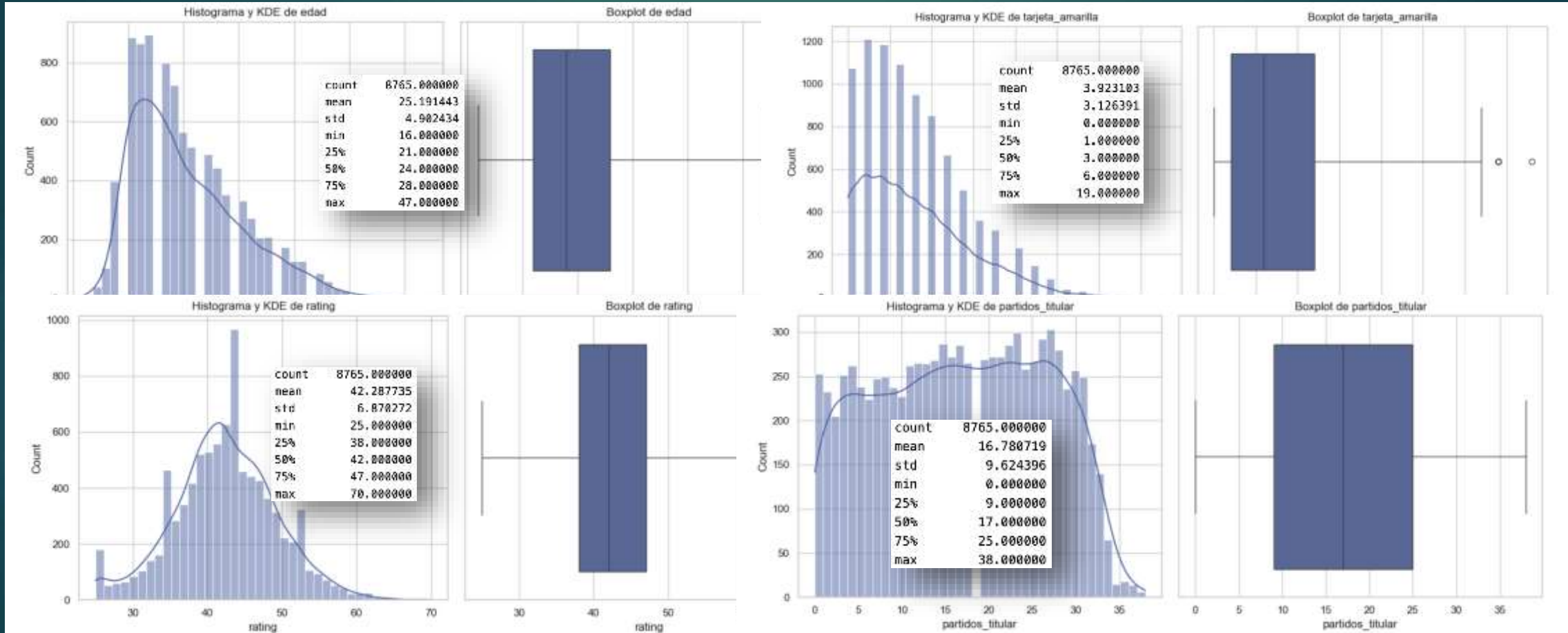
Columna/Variable	Descripción	Tipo	Cat/Num
Nombre	Nombre del jugador	String	Numerica Continua
partidos_jugados	Cantidad de partidos jugados en la temporada 2023 - 2024	Int	Numerica Discreta
partidos_titular	Cantidad de partidos que el jugador entro de titular	Int	Numerica Discreta
goles	Cantidad de goles que ha marcado en la temporada	Int	Numerica Discreta
asistencias	Número de padres/hijos a bordo	Int	Numerica Discreta
tarjeta_amarilla	Tarifa pagada por el pasajero	Int	Numerica Discreta
edad	Edad del jugador	Int	Numerica Discreta
altura	Altura del jugador	Float	Numerica Continua
valor	En Euros, cuanto vale el jugador	Float	Numerica Continua
rating	Puntaje medio del jugador en la temporada	Int	Numerica Discreta
posicion	Posicion del jugador	String	Categorica
grupo_edad	Clasificacion segun edad del jugador	String	Categorica
grupo_altura	Clasificacion segun altura del jugador	String	Categorica

. Graficos Univariantes



CATEGORICAS

. Graficos univariantes



Analisis Bivariantes



- Combinaciones y técnicas utilizadas.

- Combinación de Variables Categórica – Categórica.
(Análisis de Frecuencias)

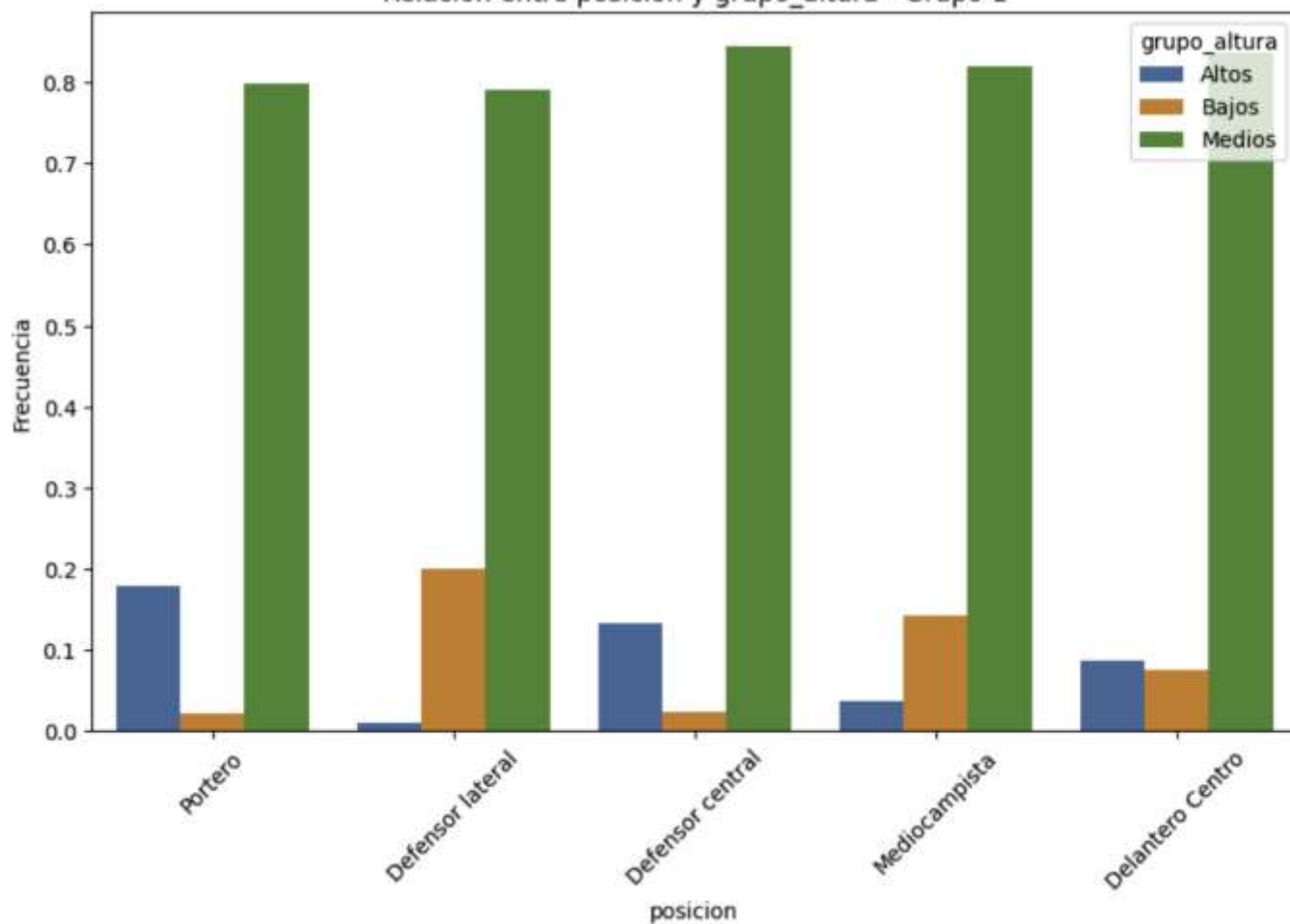
- Combinación de Variables Categóricas no binarias – Numericas.
(ANOVA).

- Combinación de variables Numérica – Numérica
(Scatterplot)

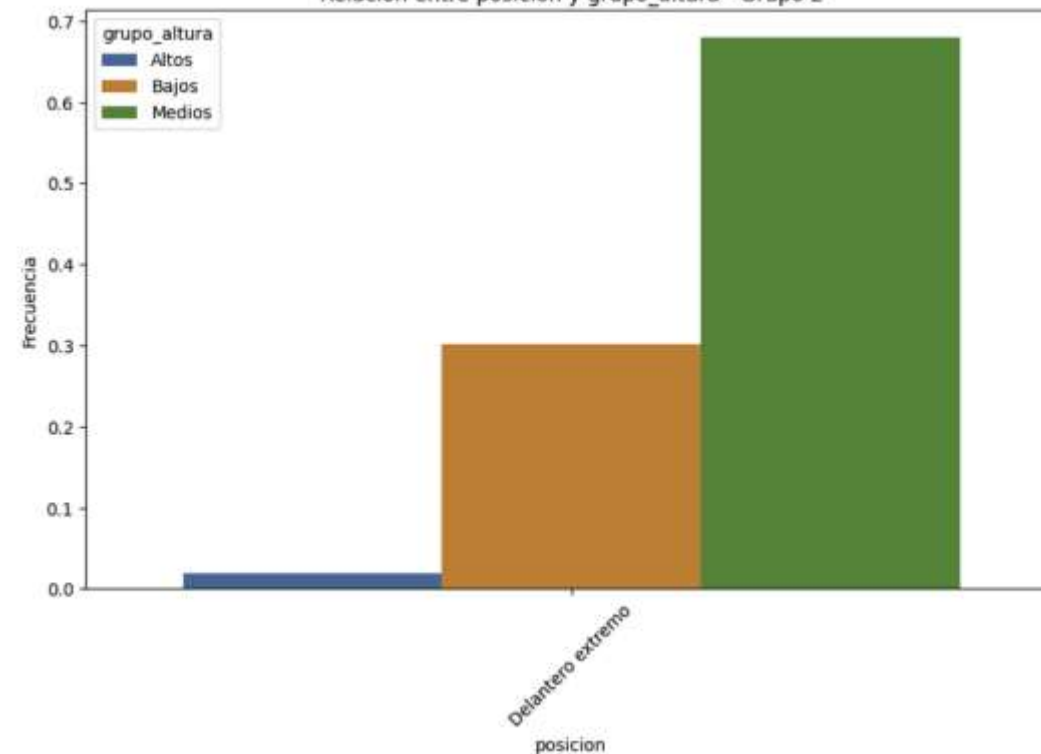
Gráficos Bivariantes



Relación entre posicion y grupo_altura - Grupo 1



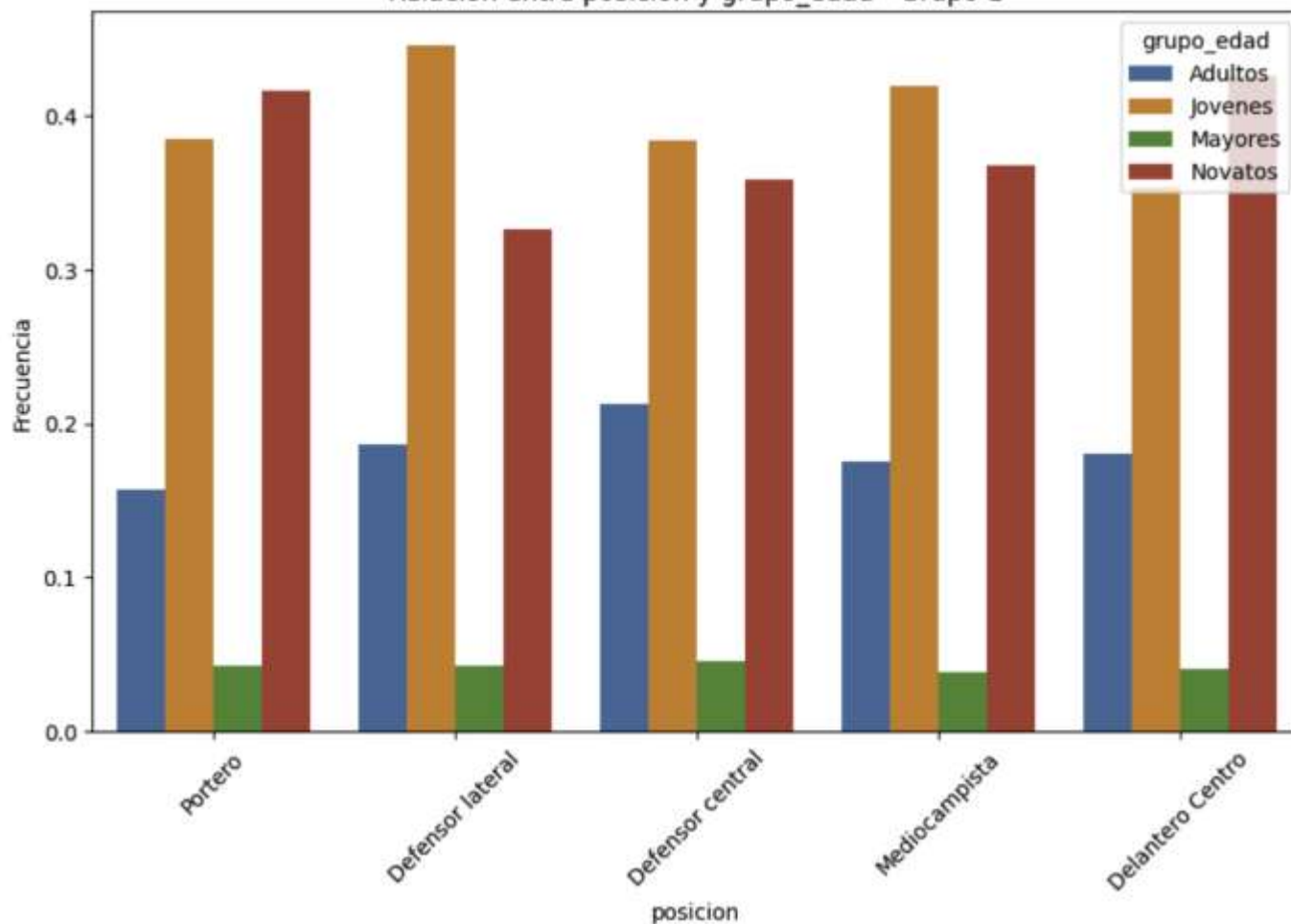
Relación entre posicion y grupo_altura - Grupo 2



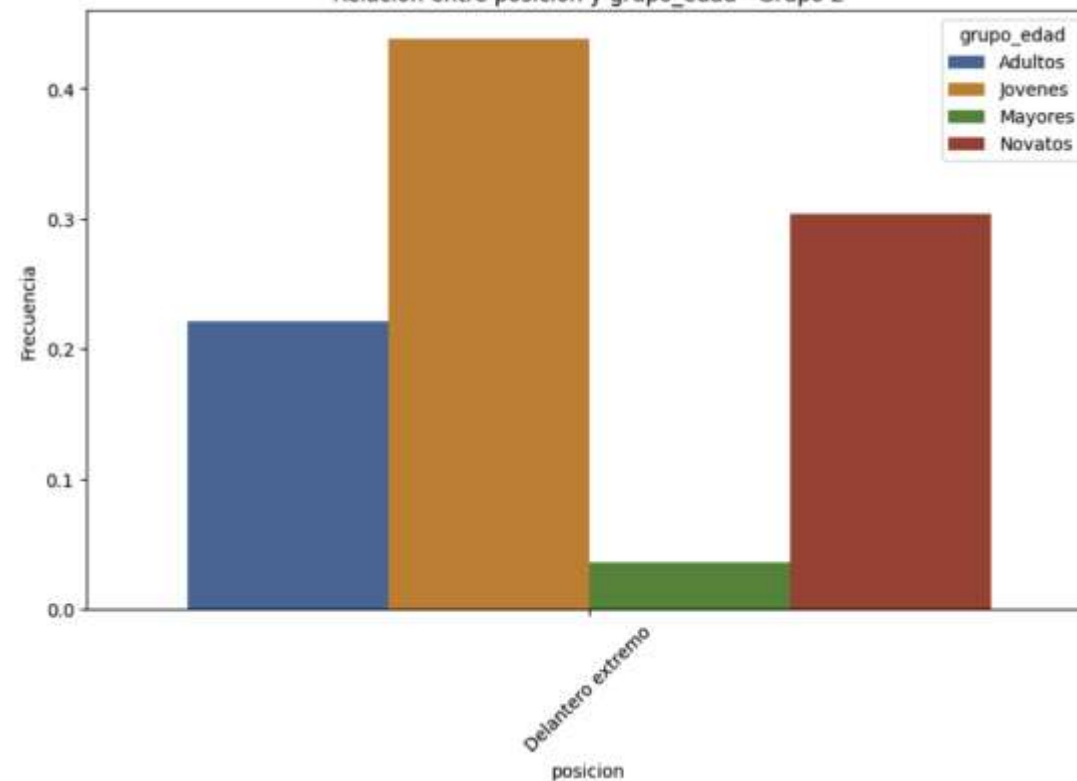
Gráficos Bivariantes



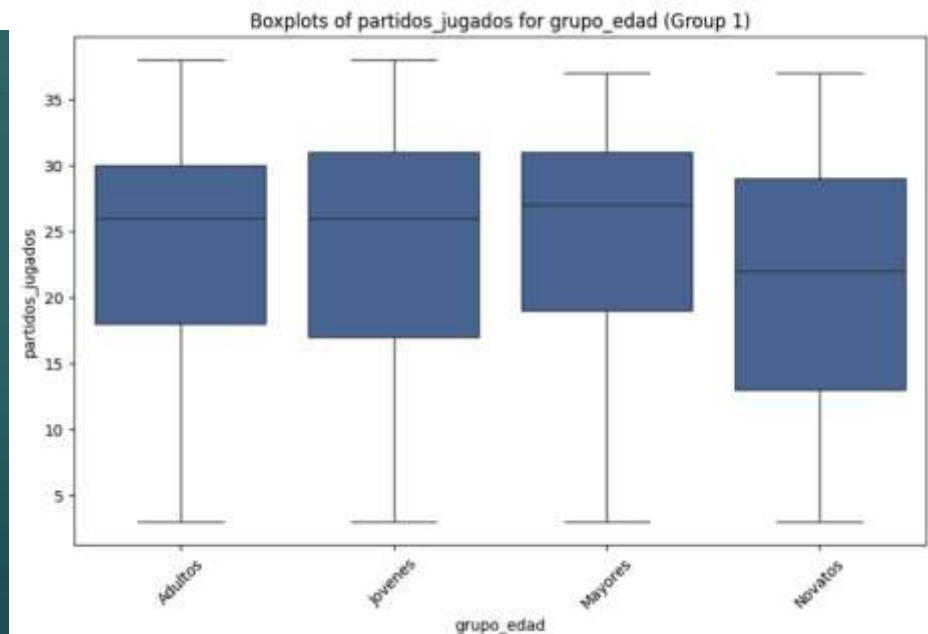
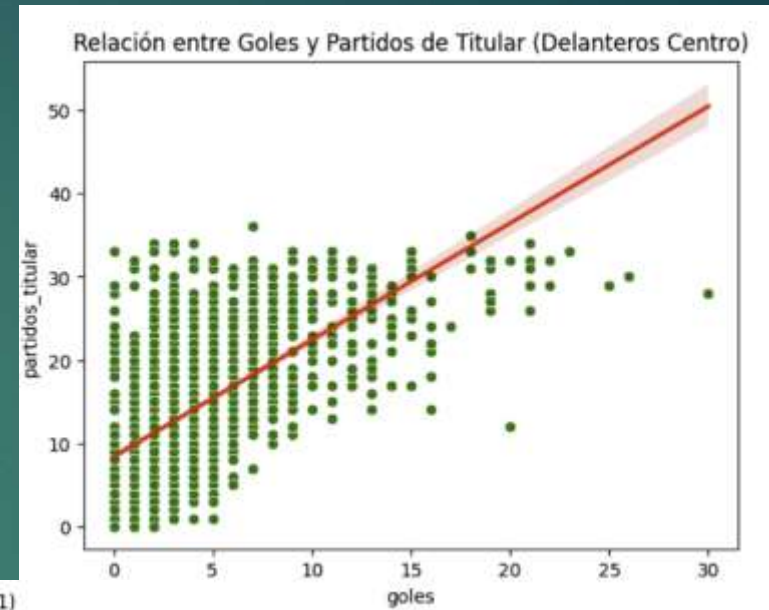
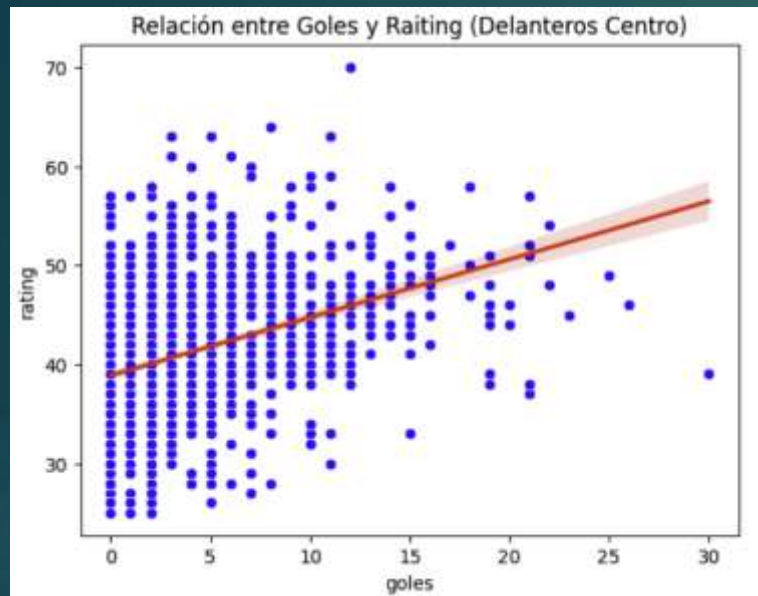
Relación entre posicion y grupo_edad - Grupo 1



Relación entre posicion y grupo_edad - Grupo 2



Gráficos Bivariantes



Conclusiones luego de análisis.



► Posición por altura y edad:

Teniendo en cuenta que es “normal” tener una estatura media ya que se lleva el 80% de valores del Data Set, entre altos (8%) y bajos (12%) se puede hacer una comparación en las posiciones.

- Si eres un jugador alto, los valores se han concentrado en las posiciones de Portero – Defensor Central – Delantero Centro. Esto era lo esperado ya que cada una de esas posiciones requiere un físico mas alto.
- Si eres jugador bajo, los valores se han concentrado en las posiciones Delantero Extremo - Defensor Lateral. Generalmente los que juegan por afuera son los jugadores que mas recorrido hacen dentro del campo de juego y los que mas rapidez tienen.

Conclusiones luego de análisis.



► Conducta por posición:

Luego del análisis, ha resultado que los jugadores detrás del medio campo y medio campo son los que mas amonestaciones tienen.

Se podía pensar que los que juegan en posiciones que no son de marca, puedan tener mas amarillas por falta de conocimiento, pero los gráficos arrojan un promedio mayor para defensores y mediocampistas, dejando en ultimo lugar al Portero (quien es el que menos interactúa en la fricción de juego) y los delanteros.

Conclusiones luego de análisis.



► Puntaje y partidos para delanteros centrales según los goles que haga:

- Al parecer un delantero centro tiene la obligación de hacer goles, ya que la relación que tienen los goles que haga para aumentar su puntaje en rating, tiene pendiente positiva.
- El mismo caso se da con los partidos que juega de titular, si hace mas goles entonces tendrá mas partidos entrando desde el principio.

Existe una diferencia en la inclinación positiva entre las pendientes que me hace pensar que tiene mas relación con los partidos que va a jugar, a que con el rating que va a obtener por los goles que el jugador hace.

Esto se da porque hay otras funciones dentro del campo que hacen el rating del jugador, no solo la de hacer goles. Eso si, el delantero que hace goles es muy probable que vaya desde el comienzo.



Conclusiones luego de análisis.

► Experiencia dentro del campo:

En este Data Set predominan los Jugadores Novatos y Jóvenes obteniendo el 76% de todos los jugadores, solo el 5% son jugadores mayores.

Viendo la cantidad de partidos jugados de los mayores a comparación de los Novatos y Jóvenes nos dimos cuenta que por mas que se lleven la mayor cantidad del data set, los entrenadores optan por tener dentro del campo a jugadores mayores que aporten su experiencia.

Su caja en el grafico me muestra que tiene una gran concentración en valores mas altos de partidos jugados y se parece al tamaño de la caja de los jóvenes.

Recomendaciones



- ▶ Si eres jugador y quieres probarte en un equipo, ten en cuenta tu altura para la posición que escojas.
- ▶ Si eres delantero centro y quieres jugar mas partidos de titular, pues haz goles. Tu raiting va a depender no solo de los goles, sino también de otros factores. (Análisis futuro).
- ▶ Si eres mayor y todavía quieres jugar profesionalmente, necesitas entrenarte para estar a la par de tus compañeros, porque la mayoría son Novatos y Jovenes.



Muchas gracias.