

2. EDA (Análisis Exploratorio de Datos)

Esta fase se centra en entender las características principales de los datos, identificar patrones, detectar anomalías y probar hipótesis iniciales, a menudo con la ayuda de visualizaciones.

Para realizar este proceso se realizó el análisis en Python y en la fase posterior se realizará un análisis a través de las visualizaciones en PowerBI que se reflejan en el Dashboard final. Todo el análisis está basado en el Athletic Club ya que es el club objetivo de este proyecto y no es necesario analizar de forma individual el resto de equipos de la liga.

En Python se realiza un análisis de las estadísticas descriptivas de los diferentes dataframes generados en fases anteriores.

Este análisis se divide en estudio de los datos globales como disparos, goles, pases, etc., un análisis de las lesiones y un análisis individual de los componentes de la plantilla.

Dentro del análisis de los datos globales se encuentran datos relevantes como;

- La edad media es más alta que la media de la liga y el ATH tiene menos internacionales que la media.
- Dominio local: En la mayoría de estadísticas consideradas como de ataque el ATH destaca por encima de la media cuando juega como local. Supera la media en remates, ataques posicionales, centros y corners entre otros. Por los datos se puede inferir que el equipo prefiere atacar por bandas ya que el número de centros es casi 1 punto superior a la media de la liga y casi 8 duelos aéreos más. Los datos también sugieren un ritmo de juego alto como muestran el alto nº de pérdidas de balón, una menor duración de las acciones y más de 7 pases hacia delante respecto a la media de la liga.
- Repliegue jugando de visitante: Los datos como visitante también muestran un equipo dominador ya que es superior a la media en muchas estadísticas pero los datos reflejan un cambio de estilo respecto a ser local. El análisis de los datos sugiere que, jugando como visitante, el ATH trata de hacer un juego algo más directo que se refleja en la diferencia en ataques posicionales de 3 respecto a su media como local. Otro aspecto destacable es que como visitante recupera más balones en el campo rival y completa menos pases.

En el análisis de lesiones se puede observar que el número total de lesiones fue de 14, teniendo una media inferior a la liga en todos los tipos de lesión y concentrado en un total de 8 jugadores, siendo Iñaki Williams el jugador con más lesiones del equipo.

El análisis individual de cada futbolista se realizará en powerBI ya que para analizar mapas de calor, tiros, posicionamiento, etc. es más útil desde esta herramienta.