

# **FIEAFURBO!**

**Análisis de datos aplicado a los videojuegos**

—

**Proyecto I, Comprensión de Datos**

Curso 2021/2022

**Grupo 1CD2 - G5**

Lluís Pellicer, Pablo Rallo, Adrián Rico, Juan Tomás y Marc Vicedo

# ÍNDICE

<b>Preámbulo</b>	<b>5</b>
Presentación	5
Agradecimientos	5
Resumen. Estado del arte	5
Palabras clave	5
<b>Introducción</b>	<b>6</b>
Preguntas iniciales	7
Objetivos	7
<b>Conjuntos de datos</b>	<b>8</b>
<b>Alcance del proyecto</b>	<b>9</b>
<b>Calendarización</b>	<b>10</b>
Actividades	10
Secuenciación. Asignación de recursos	10
Cronograma	12
<b>Análisis</b>	<b>13</b>
Análisis A. Evolución de los jugadores	13
Análisis B. Rendimiento de los equipos	14
Análisis C. Correlación Presupuesto-Rendimiento	16
<b>Resultados</b>	<b>18</b>
Estudio de los análisis	18
Producto final	22
<b>Conclusiones</b>	<b>29</b>
<b>Anexos</b>	<b>30</b>

## ÍNDICE DE FIGURAS

<b>FIGURA 1.</b> Portada del FIFA 22	<b>5</b>
<b>FIGURA 2.</b> Diagrama de precedencias del proyecto	<b>10</b>
<b>FIGURA 3.</b> Diagrama de Gantt del proyecto	<b>11</b>
<b>FIGURA 4.</b> Grafo del rendimiento a partir del Page Rank de los principales equipos europeos	<b>13</b>
<b>FIGURA 5.</b> Evolución de la valoración en el FIFA de dos jugadores	<b>17</b>
<b>FIGURA 6.</b> Page Rank de los principales equipos europeos	<b>18</b>
<b>FIGURA 7.</b> Gráfico de correlación entre el Page Rank y la valoración de los principales equipos europeos	<b>18</b>
<b>FIGURA 8.</b> Gráfico de correlación entre la puntuación en liga y la valoración de los principales equipos europeos	<b>19</b>
<b>FIGURA 9.</b> Gráfico de correlación entre el balance de gasto en jugadores y la valoración de los principales equipos europeos	<b>20</b>
<b>FIGURA 10.</b> Matriz de correlación incluyendo todas las variables estudiadas en los principales equipos europeos	<b>20</b>
<b>FIGURA 11.</b> Portal principal de la página web	<b>21</b>
<b>FIGURAS 12 Y 13.</b> Sección “Evolución” de la página web	<b>22</b>
<b>FIGURAS 14 Y 15.</b> Sección “Equipos” de la página web	<b>23</b>
<b>FIGURAS 16 Y 17.</b> Sección “Posición” de la página web	<b>24</b>
<b>FIGURA 18.</b> Sección “Blog” de la página web	<b>25</b>
<b>FIGURA 19.</b> Sección “Contáctanos” de la página web	<b>25</b>
<b>FIGURAS 20 Y 21.</b> Capturas de la cuenta de Instagram del proyecto	<b>26</b>
<b>FIGURAS 21 Y 22.</b> Capturas de la cuenta de Twitter del proyecto	<b>27</b>

## Preámbulo

### Presentación

Este trabajo ha sido realizado durante las clases de Proyecto I, del grado en Ciencia de Datos, todos los miércoles y jueves. El trabajo en clase se ha complementado con el trabajo extracurricular de todos los miembros del grupo.

A lo largo del cuatrimestre hemos analizado diversos conjuntos de datos referentes a los jugadores, partidos y equipos, tanto de la vida real como del FIFA, para observar finalmente si existe alguna relación entre el FIFA, características de los equipos y sus resultados en la vida real.

### Agradecimientos

Agradecemos principalmente a Alberto por el seguimiento que ha realizado de nuestro proyecto y todos los consejos que nos ha aportado a lo largo del cuatrimestre. Además, queremos agradecer a los profesores de otras asignaturas del primer cuatrimestre como Elena (Análisis exploratorio de datos) o Cristina (Matemática discreta), que nos han aportado conocimientos muy útiles para la realización de este trabajo.

### Resumen. Estado del arte

En el contexto actual, encontramos muchas páginas web y análisis dedicados al FIFA a sus jugadores y a los equipos del juego, pero muy pocos o prácticamente ninguno relacionan las características del FIFA con los jugadores y equipos en la vida real; por tanto, este proyecto busca establecer dicha relación para mejorar la experiencia de juego del usuario.

### Palabras clave

Fútbol / FIFA / Análisis / Correlación / Gráficos / Jugadores / Equipos / Evolución

## Introducción

Nuestro proyecto abarca la (posible) relación existente entre los datos tanto de equipos y jugadores en la vida real, con estos mismos aspectos dentro de los videojuegos de la saga FIFA.

Los juegos de la empresa Electronic Arts incluyen cada año una **valoración de cada uno de los jugadores**, basada en seis características propias del deporte como son el tiro, pase, ritmo, defensa, físico y regate. A lo largo del trabajo veremos con algunos ejemplos si la evolución de estas características tiene algo que ver con el desempeño de los jugadores en la vida real.

A continuación, **los equipos también obtienen una puntuación**, según si sus jugadores tienen mayor o menor valoración. Los equipos pueden obtener desde cinco estrellas hasta ninguna. Esto mismo también será comparado con aspectos de la vida real como el dinero invertido por parte de los clubes a lo largo de los años en fichajes, o los resultados deportivos de dichos equipos en la vida real.



Una vez realizadas estas comparaciones, seremos capaces de observar con bastante claridad **si existe alguna relación entre el FIFA y el fútbol real**. Esto lo haremos mediante técnicas de comprensión de datos, estudiadas a lo largo del curso; analizando diversos conjuntos de datos, obtenidos de diferentes fuentes o realizados por nosotros mismos (ver [Conjuntos de datos](#)), a partir de herramientas estadísticas como Statgraphics, herramientas de interfaz visual como Wordpress o Streamlit, y las librerías Pandas, Seaborg y Matplotlib (de Python), que nos proporcionarán datos y gráficas con el objetivo de esclarecer nuestras dudas acerca de esta cuestión.

## Preguntas iniciales

- ¿La **evolución** de los jugadores en el FIFA tiene algo que ver con la realidad?
- ¿Existe alguna relación entre el **rendimiento** de los clubes y su valoración en el FIFA?
- ¿El **gasto neto** en fichajes de los equipos se refleja en el desempeño de los equipos en las competiciones en las que participan?

## Objetivos

1. Comparar las valoraciones de los jugadores con su rendimiento en la vida real
2. Comparar los resultados de los equipos en la vida real con los resultados que supuestamente obtendrán en el FIFA según sus valoraciones
3. Comparar las inversiones de los equipos a lo largo de los años con el nivel del equipo en el FIFA
4. Proporcionar una página web que se pueda utilizar para visualizar el rendimiento de jugadores y equipos en el FIFA

## Conjuntos de datos

Para realizar nuestros análisis hemos utilizado cinco bases de datos, de las cuales dos ya existen y otras tres han sido creadas por nosotros:

### 1.- Bases de datos ya existentes:

1.1.-Base de datos de las **estadísticas de cada jugador** dentro del FIFA, únicamente de la temporada 21/22.

1.2.-Base de datos para el **resto de las estadísticas**, desde la temporada 14/15 hasta la 21/22.

Ambas bases permitirán un análisis pormenorizado de cada jugador, a demanda del usuario (ver [Análisis A](#)).

### 2.-Bases de datos creadas:

2.1.-Base de datos para crear un grafo del **rendimiento de los clubes** de las cuatro mejores ligas según el coeficiente UEFA, que recoge el rendimiento de los clubes de las ligas europeas en competiciones europeas durante las últimas 5 temporadas. Se centra en la temporada 20/21, teniendo en cuenta los resultados en sus respectivas ligas y en las competiciones europeas (Champions League, Europa League; la Conference League no entraría en vigor hasta la temporada siguiente). Distinguimos dos subconjuntos de datos: el 2.1.1, que refleja los **nodos** (los equipos), y el 2.1.2, que refleja las **aristas** (los enfrentamientos entre ellos), ya que trabajaremos esta base a partir de la teoría de grafos (ver [Análisis B](#)).

2.2.-Base de datos del **gasto neto en fichajes** de los equipos de la Premier League y de LaLiga desde la temporada 14/15 hasta la temporada 21/22 separado por año y el balance unificado, además de la inversión total en estos siete años.

2.3.-Base de datos de los **equipos de la Premier y LaLiga** durante la temporada 20/21 con la valoración del equipo en el juego realizada con la media del 60% de sus mejores jugadores, el Page Rank del grafo del rendimiento en esta temporada, el gasto neto en fichajes y la puntuación en la liga de dicha temporada.

Estas dos últimas bases nos permitirán establecer una correlación entre el gasto que realiza un equipo y su desempeño en las competiciones futbolísticas a lo largo de la temporada (ver [Análisis C](#)).

## Alcance del proyecto

Para que la experiencia de nuestros internautas pueda ser lo mejor posible, nuestro proyecto debe cumplir ciertas **características**. Debe ser intuitivo y fácil de utilizar, y al mismo tiempo poder analizar la evolución de los equipos y sus gastos, viendo si estos han repercutido tanto en su desempeño deportivo en temporadas posteriores como en su valoración en el propio videojuego.

Antes de evaluar el proyecto, no podemos dejar de lado también las **restricciones** que lo condicionan, ya sean desde una perspectiva económica (ya que el presupuesto con el que contamos es muy reducido), logística (la escasez de bases de datos sobre el modo online de FIFA), o deportiva (la inmensa cantidad de variables de las que depende el rendimiento individual de los jugadores, haciendo muy difícil aplicar modelos predictivos precisos sobre cada uno de ellos).

El **producto** que vamos a ofrecer a partir de este proyecto consistirá en una aplicación online accesible a todos los públicos (ver [Producto final](#)). La facción principal será una página web que contenga informes y análisis de los jugadores, relacionados con los objetivos. La propia memoria que estamos realizando, como también todos los análisis realizados a lo largo del proyecto, se verán reflejados en esta. Asimismo, crearemos diversas cuentas en redes sociales, tales como Instagram y Twitter, que nos permitan alcanzar a un mayor número de potenciales usuarios.

Como en todo proyecto, encontramos diversos **límites** que condicionan su desarrollo:

- No se indicarán los precios de los jugadores (Mercado de fichajes, Ultimate Team...)
- No se tendrá en cuenta la evolución de las cartas especiales (IF, TOTY, TOTS...), únicamente de las cartas básicas.
- No se podrán crear plantillas propias ni comprar jugadores dentro de la propia página web.
- No se evaluarán todas las ligas, tan solo las más relevantes en relación con el coeficiente UEFA (ver [Conjuntos de datos](#)).
- No se tendrá en cuenta la valoración en estrellas de los equipos.

El **alcance** estimado de nuestro proyecto es que al menos todos los jugadores de FIFA de nuestra clase conozcan y utilicen nuestra página, para esto la página web será compatible con cualquier dispositivo facilitando así que todo el mundo tenga acceso al servicio que ofrecemos, a nivel de marketing se espera que con la propia satisfacción de los usuarios sea ellos mismo quienes den a conocer nuestra página elevando así el número de usuarios que consumen nuestro producto.

## Calendarización

### I. Actividades

Para organizar mejor el desarrollo, hemos diferenciado tres **actividades a completar**, de manera que, al realizarse las tres, lleguemos al producto final. Estas son las tres secciones:

- **Análisis exploratorio de datos** (AED): A partir de las bases de datos con las que contamos (ver [Conjuntos de datos](#)), ordenar y seleccionar los datos que nos interesen, para después aplicar técnicas estudiadas en Ciencia de Datos con el objetivo de extraer conclusiones claras.
- **Creación de la memoria**: A partir de la información que recabamos a lo largo del proyecto, redactar y plasmar en un informe todo lo necesario para entender nuestro trabajo.
- **Diseño del producto**: A partir de los gráficos y análisis obtenidos tras aplicar el AED, desarrollar una página web y una marca en redes sociales (Twitter, Instagram) con una interfaz clara, única y distinguible, que nos permita obtener los resultados de dichos análisis.

### II. Secuenciación. Asignación de recursos

Dentro de cada actividad, distinguimos diversas **tareas** que debemos realizar. Aunque las tres actividades ocurren de forma prácticamente simultánea, establecemos un orden, en el que indicamos los recursos (en nuestro caso, horas de trabajo distribuidas en sesiones de dos horas) que se le asignan a cada tarea:

#### → **Análisis exploratorio de datos**

- ◆ Encontrar bases de datos → inf
- ◆ Organización/arreglo bases de datos → 1 SESIÓN (1h)
- ◆ Creación bases de datos nuevas → 4 SESIONES (8h)
- ◆ Creación de gráficos y análisis → 3 SESIONES (6/7h)

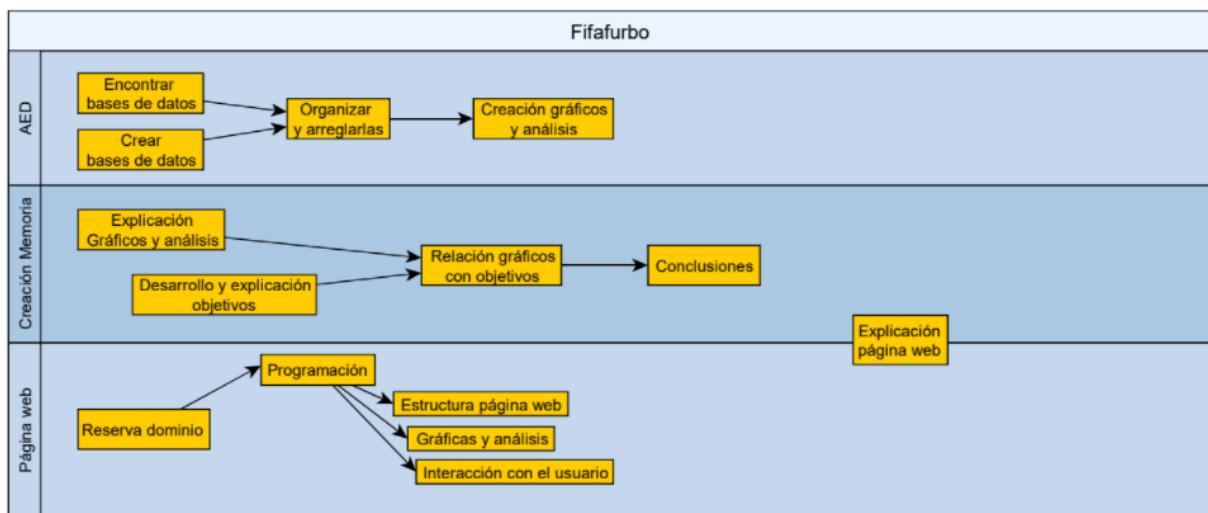
### → Creación de la memoria

- ◆ Explicación gráficos y análisis → 2 o 3 SESIONES (4h)
- ◆ Desarrollo y explicación de objetivos → 1 SESIÓN (1h 30m)
- ◆ Relación gráficos con los objetivos → 2 SESIONES (3h)
- ◆ Conclusiones → 1 SESIÓN (1h)

### → Diseño del producto

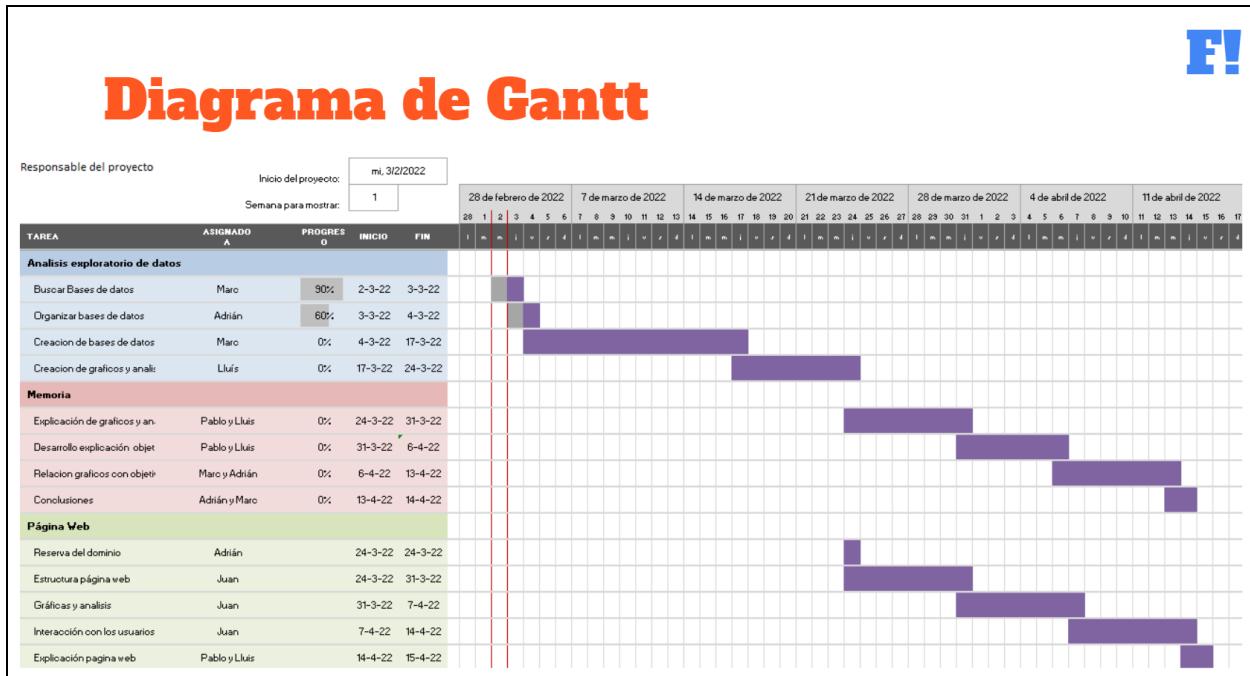
- ◆ Reserva del dominio → 1 SESIÓN (20m)
- ◆ Programación:
  - Estructura página web → 3 SESIONES (5h)
  - Gráficas y análisis → 1 o 2 SESIONES (2/3h)
  - Métodos de interacción con el usuario → 1 SESIÓN (2h)
- ◆ Explicación página web en la memoria → 1 SESIÓN (1h)

Adjuntamos un **diagrama de precedencias** que permite observar con una mayor claridad la secuenciación de las tareas:



### III. Cronograma

Para calendarizar el desarrollo de nuestras actividades y tareas, decidimos realizar un **Diagrama de Gantt**, ya que esta herramienta es muy útil a la hora de determinar aquello a lo que nos debemos dedicar a lo largo de las distintas sesiones. Como cualquier otra herramienta, este diagrama está sujeto a cambios y modificaciones causadas por diversos motivos (presentaciones intermedias del proyecto, exámenes en otras asignaturas...), por lo que estimamos oportuno adjuntar el primer prototipo de Diagrama de Gantt, realizado en el mes de febrero.



## Análisis

### Análisis A. Evolución de los jugadores

Este primer análisis, realmente, lo lleva a cabo el usuario final, ya que es él o ella quien selecciona el jugador a analizar. Utilizando Pandas y Matplotlib, dos librerías de Python, elaboramos un script que nos permite seleccionar un jugador concreto de toda la base de datos y mostrar por pantalla un gráfico de líneas en el que se distingue la valoración del jugador en cada FIFA desde el año en el que está disponible. Al depender del usuario final la elección del jugador objetivo del análisis, entendemos que él o ella conocerá la trayectoria futbolística de dicho jugador, con lo que podrá comparar la valoración a lo largo de los años con el rendimiento que haya mostrado durante las temporadas correspondientes.

La valoración que obtiene un jugador es un coeficiente entre 0 y 100, que tiene en cuenta rasgos futbolísticos: el ritmo, el tiro, el pase, los regates, el físico y las habilidades defensivas. Aunque coloquialmente se le conozca como la “media” del jugador, la valoración no se corresponde con la media aritmética de estos seis rasgos, ya que se entiende que un delantero tenga poca habilidad defensiva, o que un defensa tenga poca habilidad de tiro.

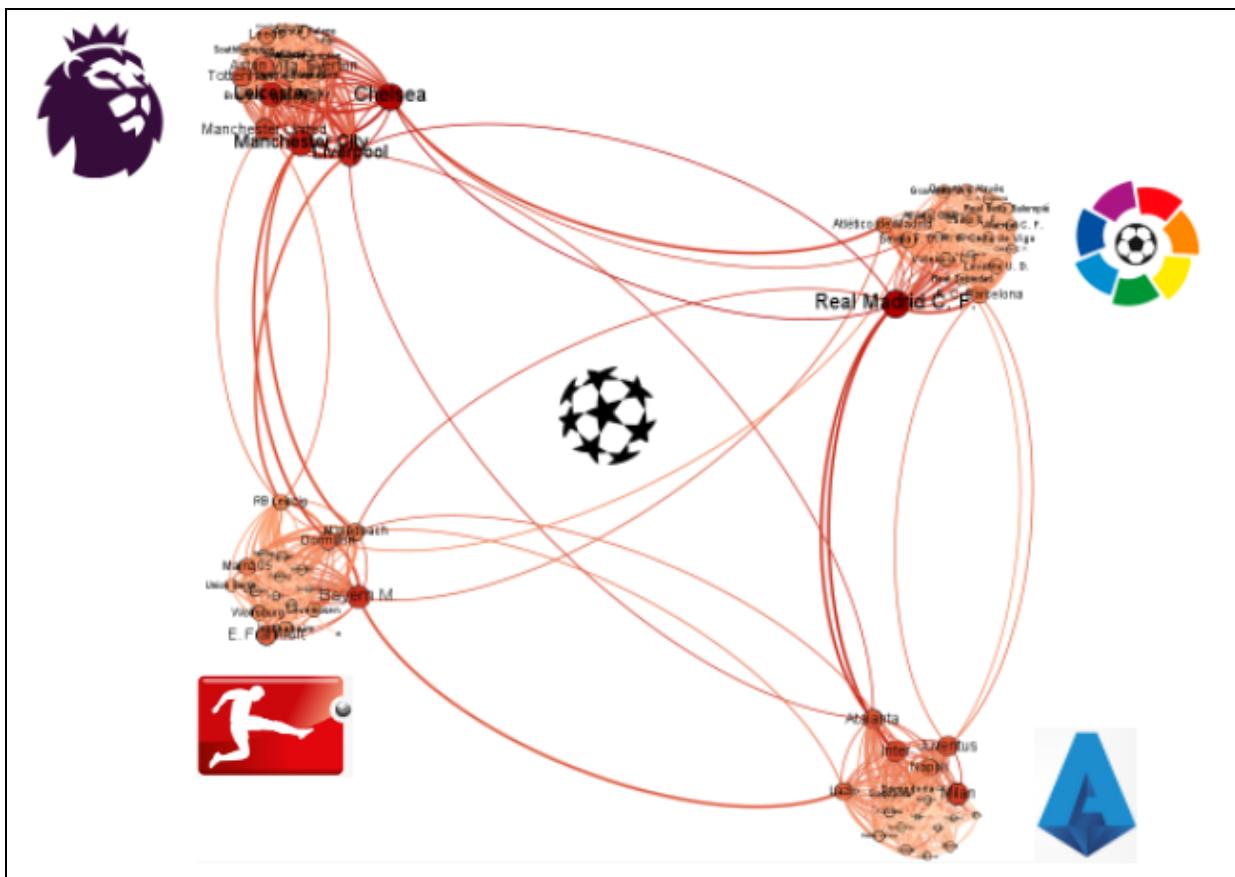
A pesar de este sistema de puntuación, muy pocos jugadores sobrepasan el coeficiente 90, ya que el juego deja algo de margen para ofrecer a lo largo de la temporada mejores cartas de cada jugador (al acabar el año se forma el TOTY o *Team of the Year*, al finalizar la temporada se realiza el mismo procedimiento de forma análoga con el TOTS o *Team of the Season...*), donde sus valoraciones suben algunos puntos. Sin embargo, como esta puntuación es algo más subjetiva y variable, dependiendo solamente de una valoración opaca y no basada en datos por parte de Electronic Arts, hemos decidido no incluirlas en los análisis, tal y como indicamos anteriormente (ver [Alcance del proyecto](#)).

Para facilitar el entendimiento de este análisis, adjuntamos dos resultados de ejemplo más adelante, para comprobar si la evolución es realista a lo largo de los años (ver [Estudio de los análisis](#)).

## Análisis B. Rendimiento de los equipos

Hemos analizado el rendimiento de todos los clubes de las cuatro mejores ligas, según el coeficiente UEFA, de la temporada 20/21. Para ello hemos realizado un grafo donde los nodos son los equipos, las aristas salen del equipo derrotado hacia el equipo victorioso y el grado son las veces que ha ocurrido en la temporada.

A continuación introducimos los datos en la aplicación Gephi, calculamos el Page Rank asociado a cada nodo y ajustamos el tamaño de estos en función de dichos resultados. Page Rank es un algoritmo, utilizado por motores de búsqueda online tales como Google, que ordena los elementos (o nodos) de un conjunto finito por orden de importancia, a partir de los enlaces existentes entre ellos (en este caso, los resultados de los partidos).



Después, hemos analizado la correlación existente entre el Page Rank y la valoración media de los clubes, calculada por nosotros realizando la media de la valoración del 60% de los mejores jugadores de cada club. Este cálculo para normalizar la variable se ha realizado de esta forma para que el programa lo tenga en cuenta plantillas de entre 15 y 20 jugadores, y no plantillas más extensas (por ejemplo, algunos equipos incluyen a canteranos, con valoraciones más bajas, en el primer equipo, lo que hace que disminuya de forma considerable el promedio de valoración). Para realizar estos análisis, de nuevo hemos hecho uso del lenguaje de programación Python; más concretamente de su librería Pandas.

Como la valoración media de los equipos depende de las valoraciones a nivel individual de sus jugadores, según la posición de defensa, mediocampo o ataque junto a la alineación habitual, posteriormente Electronic Arts asigna una puntuación en estrellas (de 1 a 5) en función de la media obtenida en cada uno de estos parámetros. Al ser una variable discreta, que solo puede tomar los valores de 1, 1.5, 2, 2.5, 3, 3.5, 4, 4.5 y 5, demasiados equipos compartirán valoración, lo cual distorsiona la percepción del verdadero potencial de un equipo. Por tanto, no nos centraremos en la valoración en estrellas sino en la media calculada por nosotros mismos (ver [Alcance del proyecto](#)).

Podemos ver cómo el equipo con mayor Page Rank no tiene por qué ser el equipo con mayor valoración media; en este caso, el Real Madrid es el equipo europeo con mayor Page Rank, pero el equipo con mayor valoración media es el Bayern de Múnich, siendo el sexto equipo según el algoritmo. Más adelante explicaremos los rasgos que inferimos a partir de esta correlación, además de incluir los primeros puestos del ranking que hemos obtenido utilizando este algoritmo (ver [Estudio de los análisis](#)).

## Análisis C. Correlación Presupuesto-Rendimiento

La finalidad de este análisis ha sido comparar el desempeño de los equipos en sus respectivas ligas y competiciones europeas, con la inversión realizada en el mercado de fichajes de la campaña anterior para comprobar si los equipos que más gastan son los que mejor rendimiento tienen.

Para ello hemos realizado una hoja de cálculo con el balance de fichajes de cada temporada desde la 2014/2015 hasta la 2021/22 tanto de la Premier League como de la Liga Española y hemos documentado detalladamente el desempeño de los 6 mejores equipos de cada temporada, quien ha sido el campeón, cómo quedó cada equipo en Champions o Europa League según el caso, quién fue el campeón de copa y qué equipos descendieron.

Posteriormente, de forma general, hemos recogido todos los galardones de cada equipo en este periodo referido en el párrafo anterior, pudiendo comprobar cuáles han sido los que más títulos han conseguido. Toda esta información luego ha sido utilizada para responder a las cuestiones planteadas, junto con los otros dos análisis (ver [Conclusiones](#)). De manera ilustrativa, exponemos los resultados correspondientes a la temporada 2020/2021 y el balance general en la liga inglesa.

### → Premier League: Temporada 2020/2021

*El Manchester City se proclamó campeón de la Premier League 2020/2021 con un desembolso de 95 millones. Participó en la Champions League, donde alcanzó la final, perdiendo contra el Chelsea. Además, ganó la Carabao Cup.*

*El segundo clasificado fue el Manchester United, con un gasto de 64 millones. Participó en la Champions League, donde cayó a la Europa League en fase de grupos. Posteriormente, perdió la final de dicho torneo contra el Villarreal.*

*El tercero fue el Liverpool con un gasto de 65 millones. Participó en la Champions League, donde quedó eliminado en cuartos de final.*

*En cuarto lugar quedó el Chelsea con un gasto de 190 millones. Participando en Champions League, donde levantó el título de campeón (también ganó la Supercopa de Europa, contra el Villarreal, aunque este partido se considera, a todos los efectos, de la siguiente temporada).*

*El quinto fue el Leicester con un gasto de 63 millones. Participó en la Europa League, cayendo en dieciseisavos de final. Por otra parte, consiguió levantar la FA Cup.*

En sexto lugar finalizó el West Ham con un gasto de 70 millones. No disputó competiciones europeas, en FA Cup fue eliminado por el Manchester United en octavos de final y en la Carabao Cup llegaron a cuartos, donde fueron eliminados por el Tottenham.

Los equipos descendidos esta Temporada fueron el Fulham, el West Bromwich Albion y el Sheffield United. Curiosamente, estos equipos no fueron de los últimos en el ranking de gasto neto, sino que ocuparon respectivamente las posiciones 12<sup>a</sup>, 13<sup>a</sup> y 10<sup>a</sup>.

## → Premier League: Resumen desde 2014

El Manchester City ha sido el equipo más laureado de Inglaterra consiguiendo 3 Premier Leagues, 1 FA Cup, 5 Carabao Cups y 2 Community Shields, con un total de 11 trofeos.

El segundo equipo con más títulos es el Chelsea con 2 Premier Leagues, 1 FA Cup, 1 Carabao Cup, 1 Champions League, 1 Europa League y 1 Supercopa de Europa, con un total también de 7 títulos, tres de ellos europeos.

El tercer equipo con más títulos es el Arsenal, con 4 Community Shield y 3 FA Cups, lo que nos deja un total de 7 trofeos.

El cuarto es el Manchester United con 1 Carabao Cup, 1 FA Cup, 1 Community Shield y 1 Europa League, lo que deja un total de 4 títulos, uno de ellos a nivel europeo.

El quinto es el Liverpool con 1 Premier League, 1 Champions League y 1 Supercopa de Europa, lo que deja un total de 3 títulos, dos de ellos europeos.

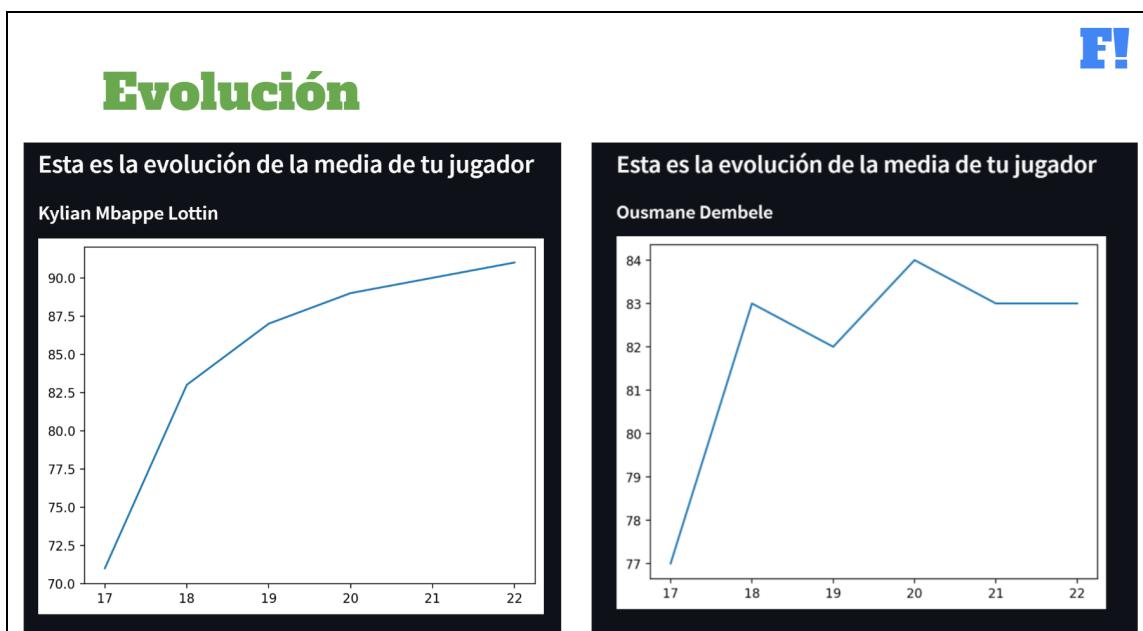
El sexto es el Leicester con 1 Premier League y 1 FA Cup, un total de dos títulos.

Por último, hemos juntado la información recogida de los presupuestos con la media que hemos calculado de los clubes en el FIFA, el Page Rank calculado anteriormente y la puntuación realizada en la temporada liguera, tanto con la Premier como con LaLiga, durante la temporada 20/21. Al reunir estos datos hemos realizado una matriz de correlación con todas estas variables, y dos gráficos x-y para poder ver si existe una correlación de los parámetros observados (ver [Estudio de los análisis](#)).

## Resultados

### Estudio de los análisis

Tal y como hemos explicado arriba, el **análisis A** no aporta resultados generales, sino específicos, ciñéndose a la solicitud de las personas usuarias. Asimismo, nosotros hemos seleccionado el caso concreto de dos futbolistas, Kylian Mbappé y Ousmane Dembelé, y tras obtener sus gráficas de evolución (ver [Análisis A](#)), hemos repasado su trayectoria futbolística en las últimas temporadas. Por ejemplo, ambos despuntaron en la temporada 16/17, correspondiente al FIFA 17, en sus respectivos equipos (Mónaco y Borussia Dortmund respectivamente), con resultados notables a nivel tanto nacional como europeo, lo cual provoca su fichaje por equipos más grandes; PSG en el caso del primero, Barcelona en el caso del segundo. Esto hace que la valoración de ambos suba en el FIFA 18. A pesar de ello, cabe destacar que Dembelé sufrió diversas lesiones de larga duración en el transcurso de la temporada 17/18, de forma que su rendimiento bajó de forma considerable. Esto hizo que su valoración se resintiera en el FIFA 19. De la misma forma que Dembelé no llegó a destacar en el Barcelona a su máximo nivel mostrado en Alemania, Mbappé ha seguido mejorando a nivel individual de forma considerable, por lo que su valoración ha seguido una tendencia ascendente en su trayectoria en el PSG.



Por otra parte, a partir del grafo visto en el **análisis B**, con la aplicación Gephi hemos obtenido el Page Rank de cada uno de los nodos; es decir, el rendimiento de cada equipo en la temporada medida en el número de victorias y la importancia de estas.

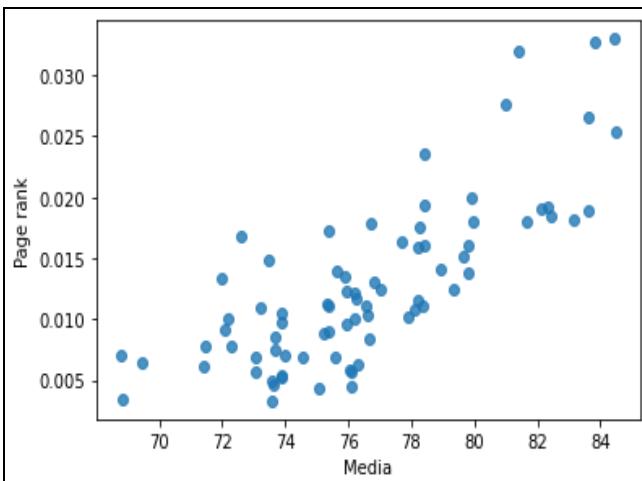
En esta tabla se pueden observar los clubes de las cuatro grandes ligas europeas que tuvieron un mejor rendimiento durante la temporada 20/21, teniendo en cuenta el Page Rank.

Se observa que todos ellos son clubes con un buen rendimiento en sus ligas o que han llegado lejos en competiciones europeas. A pesar de esto, cabe destacar que este algoritmo tiende a “sobredimensionar” a los equipos ingleses (por ejemplo, Everton o Arsenal por delante de Atlético o Barcelona), ya que considera más importantes los partidos de la Premier que los de la liga española; al ser en los últimos años una liga tan igualada, que uno de los grandes equipos pierda contra uno inferior le concede más puntos al segundo de los que pueda conceder un resultado similar en nuestra liga.

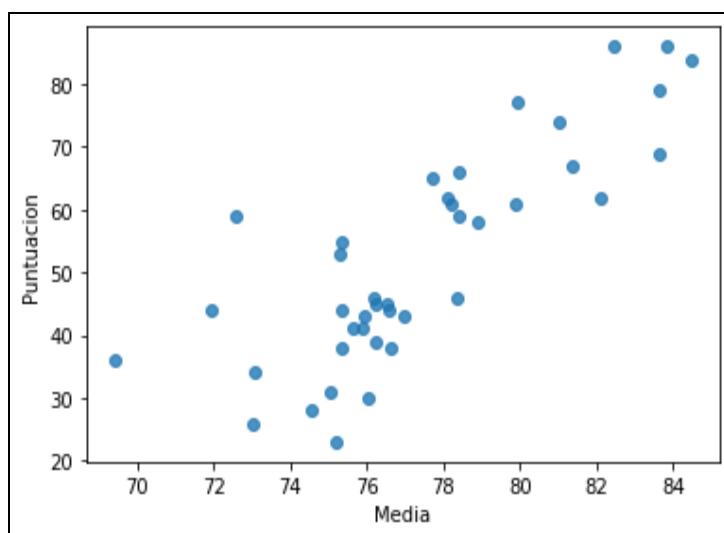
Label	pageranks
Real Madrid C. F.	0.032971
Manchester City	0.032663
Chelsea	0.032014
Manchester United	0.027539
Liverpool	0.026498
Bayern M.	0.025286
Leicester	0.023519
arsenal	0.019979
Everton	0.019313
Inter	0.019138
Tottenham	0.019045
F. C. Barcelona	0.018956
Atletico de Madrid	0.018459
Juventus	0.018196
Dormund	0.018072
Sevilla F. C.	0.018065
E. Frankfurt	0.017792

En este gráfico, adjuntado en el apartado correspondiente del análisis (ver [Análisis B](#)), se puede observar cómo el Page Rank y la media que han sido calculados anteriormente tienen

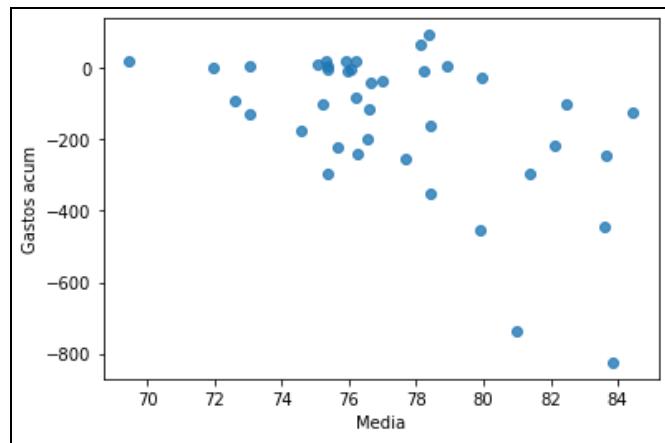
entre sí una correlación lineal directa, ya que cuánto más aumenta la media en el videojuego mejor es su rendimiento en la temporada. Como la valoración de los jugadores depende de su rendimiento, consideramos que esta relación es prácticamente fuerte, ya que supera el valor de 0.8 en el coeficiente de correlación.



Asimismo, podemos obtener resultados a partir del análisis B y del C, ya que, si establecemos un gráfico de correlación entre la valoración media calculada de cada equipo y la puntuación que obtuvo en su liga nacional durante la etapa estudiada, podemos ver que los equipos con mayor valoración media son los que más puntos obtuvieron. Sin embargo, el coeficiente de correlación no será tan elevado, ya que, como hemos dicho antes, hay una sobrerepresentación de equipos ingleses de media tabla, lo cual hace que sean equipos con menor puntuación.



El dinero no lo es todo, pero tampoco podemos malinterpretar la información y pensar que esto no va a tener ninguna repercusión, ya que si el proyecto deportivo es bueno, con la dirección deportiva y el entrenador en consonancia, una mayor inversión y por lo tanto una adquisición de mejores jugadores que vengan a cumplir un rol determinado sí que sería de gran ayuda y podría aumentar tanto el rendimiento en el corto plazo, como posteriormente el desempeño en la temporada y en la consecución de títulos a largo plazo. La gráfica del lateral nos muestra en el eje X la valoración media, y en el eje Y el gasto realizado por los equipos. La correlación es inversa (negativa), ya que los valores reflejan pérdidas; habiendo incluido únicamente el balance de fichajes, no se tiene en cuenta el gasto en dos facetas económicas fundamentales en el fútbol moderno: la publicidad y los derechos de televisión. Este último es mucho más destacable en la Premier League.



Aquí tenemos la matriz de correlaciones que hemos estado mencionando anteriormente y de la cual hemos obtenido los valores del coeficiente entre cinco valores de referencia (valoración media del equipo, Page Rank, gasto en la temporada 20/21, resultado en la temporada 20/21 y gasto acumulado desde 2014) que posteriormente hemos utilizado para llegar a las conclusiones expuestas.

	overall	page_rank	gastos	puntuacion	gastos_acum
overall	1	0.78	-0.096	0.81	-0.53
page_rank	0.78	1	-0.36	0.84	-0.67
gastos	-0.096	-0.36	1	-0.14	0.45
puntuacion	0.81	0.84	-0.14	1	-0.52
gastos_acum	-0.53	-0.67	0.45	-0.52	1

## Producto final

Para que todas las personas que así lo deseen puedan consultar el resultado de los análisis expuestos con anterioridad, consideramos necesario implementar una plataforma online que lo permita. Es precisamente por esta razón por la que hemos creado una **página web** y dos **cuentas en redes sociales** (Twitter e Instagram).

Por una parte, en la página web ([enlace](#)), realizada con Wordpress, utilizamos una interfaz minimalista, con la que buscamos que destaque las imágenes. Incluimos en ella diversos enlaces a aplicaciones Streamlit que reflejan los análisis, ya que este servicio nos permite ejecutar scripts de Python en línea. Los scripts están subidos a un repositorio de Github; asimismo, los adjuntamos para que puedan ser consultados. Dentro de la página distinguimos diferentes secciones:

- La Página principal, en la que exponemos de forma breve en qué consiste nuestro proyecto.

**F!**  
**FIFAFURBO!**

Bienvenido al siguiente nivel.

[Evolución](#)  
[Equipos](#)  
[Posición](#)  
[Blog](#)  
[Contáctanos](#)

## Análisis de datos aplicado a los juegos

Somos un grupo de estudiantes del Grado en Ciencia de Datos de la Universitat Politècnica de València, que está realizando un proyecto de larga duración relacionado con la **comprensión de datos**. Desde el profesorado se nos proponía aplicar los conocimientos que teníamos en la materia a situaciones reales, que estuvieran respaldadas por bases de datos lo suficientemente amplias como para ser analizadas. Nosotros decidimos fijarnos en el FIFA, ya que nos ofrece muchas variables a estudiar.



Escola Tècnica Superior d'Enginyeria Informàtica  etsinf

UNIVERSITAT POLITÈCNICA DE VALÈNCIA

- Evolución, donde, tras seleccionar a un jugador, la interfaz nos muestra con un gráfico de líneas la media de valoración que tiene dicho jugador en el FIFA a lo largo de los años.

**F!**  
**FIFAFURBO!**

Bienvenido al siguiente nivel.

Evolución  
Equipos  
Posición  
Blog  
Contáctanos

**Evolución**  
Desde sus inicios, hasta hoy



Prácticamente cada día, aparecen nuevos jugadores con un gran potencial. El tipo de jugador que «sabes» que, en unos años, estará entre los mejores. En esta página podrás consultar la evolución en el FIFA del jugador que decidas a lo largo de los últimos 8 años.

Pulsa aquí



Aquí podrás analizar la evolución del jugador que prefieras.

¿A quién buscas? (mínimo 3 caracteres)

Por defecto siempre se selecciona el primero de la lista, puedes cambiar el número cuando quieras

- Equipos, que nos permite consultar la valoración de los equipos en las grandes ligas europeas, según la fórmula aplicada a la valoración de los jugadores (ver Análisis B) desde el FIFA 15 al 22; es decir, podemos consultar un *power ranking* de las grandes ligas desde la temporada 14/15 hasta la 21/22 realizado al comienzo de las mismas.

**F!**  
**FIFAFURBO!**

Bienvenido al siguiente nivel.

[Evolución](#)  
[Equipos](#)  
[Posición](#)  
[Blog](#)  
[Contáctanos](#)

Twitter icon | Instagram icon

Equipo

¿Quiénes se llevarán los títulos?

WINNERS  
UEFA EUROPA LEAGUE 2020/21

Pulsa aquí

**FIFAFURBO!**

¿Qué liga quieres comprobar?  
English Premier League

¿Qué FIFA quieres comprobar?  
15

Valoración media clubs English Premier League fifa 15

Club	Valoración Media
Chelsea	~8.5
Manchester City	~8.2

- Posición, donde podemos filtrar los datos de los jugadores de la base de datos para así obtener los diez jugadores con mejor valoración que cumplan los requisitos que se hayan establecido: nacionalidad, posición o liga.

The screenshot shows the 'Posición' (Position) screen of the FIFAFURBO! app. At the top, there's a large red 'F!' logo followed by the word 'FIFAFURBO!' in a bold, white, sans-serif font. Below the logo, the text 'Encuentra al mejor en cada puesto' (Find the best in every position) is displayed. On the left side, there's a sidebar with links: 'Bienvenido al siguiente nivel.', 'Evolución', 'Equipos', 'Posición', 'Blog', and 'Contáctanos'. Below these links are social media icons for Twitter and Instagram. In the center, there are two small images: one of a coach pointing on a soccer field and another of a man in a suit talking to another person. To the right of the images, the text reads: 'Aquí podrás seleccionar a los mejores jugadores de cada posición, según las estadísticas del FIFA 22. La aplicación utiliza las abreviaturas inglesas para cada posición, y un código que asigna un número a cada nacionalidad y a cada liga. En estas tablas puedes ver lo que tienes que escribir:' followed by a red button labeled 'Pulsa aquí' (Press here).

The screenshot shows the main search interface of the FIFAFURBO! app. At the top, the 'FIFAFURBO!' logo is centered between two diagonal lines. Below the logo, the text 'Aquí podrás filtrar los mejores jugadores por liga, nacionalidad o posición.' (Here you can filter the best players by league, nationality or position.) is displayed. There are three input fields with placeholder text: 'Elige una nacionalidad (escribe el número)' (Choose a nationality (write the number)), 'Elige una posición (escribe la abreviatura inglesa: por ejemplo, portero es GK)' (Choose a position (write the English abbreviation: for example, goalkeeper is GK)), and 'Elige una liga (escribe el número)' (Choose a league (write the number)). Below these fields, a message says 'Estos son los mejores jugadores que cumplen tus requisitos:' (These are the best players who meet your requirements:). The entire interface has a dark background with light-colored text and input fields.

- Blog, que incluye entradas informativas sobre nuestro proyecto: en la actualidad solamente hay una, del día que hicimos pública la página.

**F!**

## FIFAFURBO!

Bienvenido al siguiente nivel.

[Evolución](#)  
[Equipos](#)  
[Posición](#)  
[Blog](#)  
[Contáctanos](#)

### Nace FIFAFURBO!



Presentamos un proyecto que pretende ayudar a los jugadores de FIFA a comprender mejor el juego, y así puedan tomar mejores decisiones en base a nuestro análisis de datos: podemos ver la evolución de un jugador en los últimos años, comparar la valoración de un equipo en el FIFA y sus resultados en las ligas [...]

[23 de marzo de 2022](#)

- Contáctanos, donde incluimos información sobre la ETSINF.

**F!**

## FIFAFURBO!

Bienvenido al siguiente nivel.

[Evolución](#)  
[Equipos](#)  
[Posición](#)  
[Blog](#)  
[Contáctanos](#)

### Contáctanos

**Escola Tècnica Superior  
d'Enginyeria Informàtica**

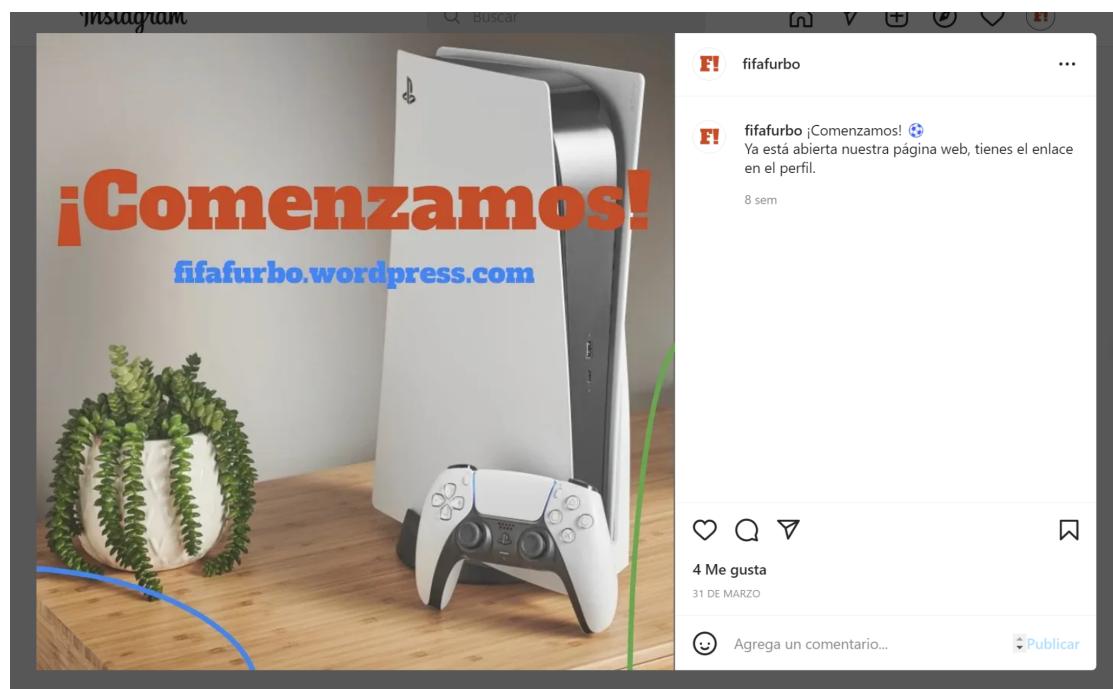


Edificios 1E – 1G – 1H

Camí de Vera, s/n, 46022 València, España

Por otra parte, las cuentas en redes sociales, accesibles desde la página web, también contendrán información, de manera más breve, sobre lo que aparece en los análisis. En ambas biografías, aparecen enlaces a la página web.

→ Instagram ([enlace](#)):



→ Twitter ([enlace](#)):



**FIFAFURBO!**  
@fifafurbo\_

¿Quieres saber cómo ha evolucionado tu jugador favorito en los últimos años? Ahora puedes

[fifafurbo.wordpress.com](https://fifafurbo.wordpress.com)

**Análisis de datos aplicado a los juegos**

Somos un grupo de estudiantes del Grado en Ciencia de Datos de la Universitat Politècnica de València, que está realizando un proyecto de larga duración relacionado con la **comprensión de datos**. Desde el profesorado se nos proponía aplicar los conocimientos que teníamos en la materia a situaciones reales, que estuvieran respaldadas por bases de datos lo suficientemente amplias como para ser analizadas. Nosotros decidimos fijarnos en el FIFA, ya que nos ofrece muchas variables a estudiar.

Evolución  
Equipos  
Posición  
Blog  
Contáctanos

0:28 | 16 reproducciones

8:52 a. m. · 11 may. 2022 · Twitter Web App

## Conclusiones

Después de todo el trabajo realizado, la búsqueda de datos, tratamiento de ellos y los análisis realizados con la finalidad de responder a las preguntas mencionadas con anterioridad (ver [Preguntas iniciales](#)), podemos ofrecer respuestas con seguridad a las cuestiones que nos hemos planteado.

Las dos primeras preguntas tienen una respuesta clara: **el rendimiento de jugadores y equipos durante la temporada sí que tiene relación con las valoraciones obtenidas**. Tanto los jugadores de forma individual, como los clubes, evolucionan a lo largo de los años, y adquieren mayor relevancia con el paso del tiempo los análisis que hayamos podido realizar. Pese a que el análisis B es una “foto fija” de la última temporada de la que disponemos datos de forma completa, los gráficos de correlación muestran claramente cómo los equipos mejor valorados son aquellos que obtienen mejores resultados.

Respecto a la tercera pregunta, la primera conclusión que hemos sacado al comparar en la matriz de correlación (ver [Estudio de los análisis](#)) el gasto de los equipos con otras variables que se relacionen con el rendimiento de los equipos ha sido que **el gasto de un club un año en concreto normalmente no tiene relación con su actuación ese año**. Los índices de correlación que tiene el gasto anual con estas variables son valores negativos cercanos a 0; todos estos indican que la correlación que existe entre ellas es débil. Cabe destacar que los valores son negativos no porque la relación que existe sea inversa en sí, sino porque al crear las bases de datos las pérdidas se reflejan en la base de datos en números negativos.

Después de ver esto, decidimos comparar estas mismas variables con el gasto neto total acumulado de las últimas temporadas, para así tener un punto de vista más amplio y llegar a una conclusión que se corresponda más con la realidad. En este análisis, efectivamente, los índices de correlación se acercan más a -1, y podemos concluir que la correlación entre estas variables es algo más fuerte. Esta conclusión es lógica, un equipo no se construye únicamente en un año, por lo que por mucho que gaste un club en fichajes un año concreto no se asegura conseguir unos grandes resultados al final de la temporada. Sin embargo, **si la inversión económica se mantiene a lo largo del tiempo el nivel del equipo va creciendo de manera sostenida** y es mucho más probable que consiga unos resultados mejores.

## Anexos

Al ser este trabajo eminentemente de investigación y elaboración propia, no tenemos una bibliografía a la que acudir, sino que hemos consultado los resultados de las principales competiciones futbolísticas de Europa en los últimos ocho años, además de los balances de gasto disponibles de los equipos que competían en ellas. Pasamos a detallar todos los ficheros anexos a esta memoria, tanto los obtenidos mediante búsqueda como los elaborados por nosotros mismos:

- **Base de datos 1.1:** estadísticas y valoración por jugador, en el FIFA 22
- **Base de datos 1.2:** estadísticas y valoración por jugador, desde el FIFA 15 hasta el FIFA 22
- **Base de datos 2.1:** rendimiento de los clubes de las cuatro principales ligas europeas
- **Base de datos 2.2:** gasto neto en fichajes de Premier y LaLiga
- **Base de datos 2.3:** variables de rendimiento de equipos aplicables a Premier y LaLiga
- **Grafo** del Análisis B
- Aplicación de Streamlit: **evolucion.py**
- Aplicación de Streamlit: **clubs.py**
- Aplicación de Streamlit: **posicion.py**

