



Universidad de Valladolid



ESCUELA DE INGENIERÍAS
INDUSTRIALES

UNIVERSIDAD DE VALLADOLID
ESCUELA DE INGENIERIAS INDUSTRIALES

Grado en Ingeniería en Organización Industrial

HERRAMIENTA DE INTELIGENCIA DE NEGOCIO
DEPORTIVO DESARROLLADA EN POWER
BI COMO APOYO EN LA BÚSQUEDA DE LAS
MEJORES ALTERNATIVAS A LA HORA
DE REEMPLAZAR LA BAJA DE UN JUGADOR DE
FÚTBOL

Autor:

Machuca Blázquez, Sergio

Tutor:

Villafáñez Cardeñoso, Félix Antonio

Departamento de Organización de Empresas y
Comercialización e Investigación de Mercados

Valladolid, Junio de 2023.





AGRADECIMIENTOS

Quiero dedicar este proyecto a mis padres, hermano, familia y amigos. A mis tíos y a mis primos pequeños, los cuales, en estos dos últimos meses de realización del trabajo en Holanda, han sido un apoyo fundamental para mí.

Me gustaría agradecer a todos haber creído y confiado en mi en los momentos más complicados y, sobre todo, en apoyarme en las difíciles decisiones tomadas en los últimos meses, sin ellos no habría sido posible.

Se cierra la mejor etapa de mi vida, pero la que viene va a ser mucho mejor.





RESUMEN

El presente Trabajo de Fin de Grado (TFG) consiste en la planificación y realización de un proyecto de Business Intelligence Deportiva, cuyo resultado final sea un cuadro de mando que sirva como apoyo a la toma de decisiones por parte del director deportivo de un club de fútbol a la hora de reemplazar la posible baja de un jugador. Como punto de partida se plantea un caso práctico basado en equipos y jugadores reales.

El cuadro de mando desarrollado deberá proporcionar al usuario, a través de una interacción fácil e intuitiva, tanto la información procesada como aquellos conocimientos que le resulten valiosos a la hora de evaluar los mejores candidatos para reemplazar a un jugador de la plantilla, permitiendo así, además, ahorrar costes y acotar los tiempos en la toma de decisiones en este ámbito. La implementación del mismo se deberá realizar mediante el software Power BI de Microsoft.

Palabras clave:

Inteligencia de Negocio, Fútbol, Scouting, Power BI, Toma de Decisiones.





ÍNDICE DE CONTENIDO

AGRADECIMIENTOS.....	III
RESUMEN.....	V
ÍNDICE DE CONTENIDO.....	7
1. INTRODUCCIÓN Y OBJETIVOS.....	11
1.1 Introducción.....	11
1.2 Contexto y justificación del trabajo realizado.....	11
1.3 Objetivos.....	12
1.4 Descripción del trabajo planificado para alcanzar los objetivos propuestos.....	12
1.5 Motivación.....	14
2. BIG DATA.....	15
2.1 ¿Qué es el Big Data?.....	15
2.2 Historia del Big Data.....	15
2.3 Nacimiento del término Big Data.....	15
2.4 Big data en la actualidad. (Beneficios y usos).....	16
2.5 Big Data en el futuro.....	18
3. INTELIGENCIA DE NEGOCIO (BUSINESS INTELLIGENCE).....	21
3.1 ¿Qué es el Business Intelligence?.....	21
3.2 Funcionamiento de la inteligencia de negocio.....	21
3.3 Ventajas de la Inteligencia de negocio.....	22
4. BIG DATA DEPORTIVO.....	27
4.1 ¿Qué es el Sports Analytics?.....	27
4.2 Orígenes del Big Data en el deporte.....	27
4.2.1 Sabermetría.....	27
4.2.2 Moneyball.....	27
4.2.3 Muthuball.....	28
4.3 La importancia del Big data en el deporte.....	29
5. BIG DATA EN EL FÚTBOL.....	33
5.1 Usos del Sport Analytics en el mundo del fútbol.....	33
5.2 Proveedores de datos y microchips GPS.....	34
5.3 Eventing y Tracking.....	36
5.3.1 Eventing.....	36
SERGIO MACHUCA BLÁZQUEZ.....	7



5.3.2 Tracking	40
5.4 Wellness control y prevención de lesiones	42
5.4.1 Introducción.....	42
5.4.2 Parámetros clave en la prevención de lesiones	43
5.4.3 Wellness control	45
5.5 Scouting deportivo y detención del talento	46
5.5.1 Introducción.....	46
5.5.2 Datos técnico – tácticos.....	48
5.5.3 Datos físicos	49
5.5.4 Datos subjetivos	49
5.5.5 Datos cognitivos	50
5.5.6 Herramientas de Scouting Deportivo y detención del talento.....	50
6. CASO PRÁCTICO.....	53
6.1 Introducción	53
6.2 Identificación del cliente	53
6.2.1 Borussia Dortmund.....	54
6.2.2 Sebastian Kehl.....	55
6.2.3 Edin Terzic.....	57
6.3 Petición del cliente (Requerimientos).....	58
6.4 Presupuesto del Proyecto	60
6.4.1 Estimación de costes por actividad	60
6.4.2 Coste mensual	61
6.5 Proceso de ETL.....	61
6.5.1 Extracción de los datos	61
6.5.2 Transformación y carga de los datos.....	64
6.6 Definición de KPI’s o indicadores	69
6.7 Creación de algoritmo de similitud	77
6.7.1 Asignación de pesos a las variables.....	77
6.7.2 Creación de un algoritmo de similitud para cada variable.....	79
6.7.3 Creación de algoritmo de similitud total	80
6.8 Visualización y cuadro de mandos	81
6.8.1 Portada	81
6.8.2 Algoritmo de similitud y scouting de jugadores	83



6.9 Informe de jugadores	97
7. RESULTADOS OBTENIDOS Y OPINIÓN PERSONAL	101
7.1 Resultados obtenidos	101
7.2 Opinión personal	103
REFERENCIAS BIBLIOGRÁFICAS.....	105
ÍNDICE DE SIGLAS.....	109
ÍNDICE DE TABLAS.....	111
ÍNDICE DE IMÁGENES	113
ÍNDICE DE ILUSTRACIONES	115





1. INTRODUCCIÓN Y OBJETIVOS

1.1 Introducción

La Ingeniería Industrial es una de las ramas más amplias de la actualidad. Trata numerosos temas, desde cualquier tipo de problema relacionado con los números, hasta establecer las bases de un proyecto. Por ello, cuándo tan sólo tenía 17 años me embarqué en un viaje que estoy a punto de finalizar.

De pequeño siempre me gustaron los números, tampoco eran una pasión para mí, pero prefería usar la cabeza para resolver problemas que estudiar párrafos y párrafos de memoria. De hecho, siempre sacaba mejores notas en asignaturas como matemáticas y ciencias de la naturaleza antes que en asignaturas como literatura o ciencias sociales. Al llegar a segundo de bachillerato y tampoco tener muy claro que carrera elegir, decidí estudiar una ingeniería por la cantidad de salidas que esta tenía, aparte de porque como ya he comentado anteriormente, me gustaban más los números que las letras.

Por casualidades de la vida, acabé en Valladolid, ya que mi primera intención era estudiar una ingeniería en Zaragoza (ya que soy de Soria y está más cerca) pero no me admitieron en ningún Colegio Mayor.

No puedo estar más contento ni de la decisión que tomé ni de la suerte que he tenido en conocer toda la gente que me ha acompañado en estos 6 años estudiando el grado de Ingeniería en Organización Industrial en la Universidad de Valladolid.

1.2 Contexto y justificación del trabajo realizado

Desde pequeño me han encantado los deportes y sobre todo el fútbol, pero ¿Cómo relacionar esta pasión con la Ingeniería Industrial? Es algo que me llevaba causando bastante interés durante mucho tiempo y por fin encontré la manera. El mundo de los datos.

En la actualidad, a través de los avances tecnológicos, cada vez es más fácil conseguir cualquier tipo de dato en cualquier rama, desde la contabilidad de una empresa, hasta saber qué velocidad máxima puede alcanzar un jugador de fútbol (que es del tema que nos incumbe).

Además, se puede ver simplemente al ver cualquier partido de fútbol, que, a parte de las estadísticas del juego finales, Microsoft ha desarrollado una herramienta basada en los datos para poder predecir el porcentaje de posibilidad de que un equipo gane al otro, o de que un tiro se convierta en gol, en función del jugador, distancia a portería, tiros realizados a esta distancia de ese jugador... e innumerables datos que podría seguir escribiendo en hojas y hojas.

Este trabajo de fin de Carrera pretende justificar todo lo resumido en los párrafos anteriores además de enseñar cómo, a través de los datos, podemos crear una



herramienta de negocio deportivo para encontrar el tipo de jugador que mejor encaje en nuestro equipo en función de diferentes factores.

1.3 Objetivos

El objetivo principal de este TFG es realizar un proyecto de Business Intelligence cuyo resultado final es un cuadro de mando, fácil e intuitivo, para ahorrar y acotar tiempo a nuestro director deportivo (persona encargada de tomar las decisiones en el ámbito deportivo de un club) a la hora de realizar los diferentes fichajes de la entidad. La implementación de este se realiza a través del software Power Bi de Microsoft.

Además, podemos destacar una serie de objetivos secundarios, en los cuales profundizaremos más en detalle a lo largo del presente trabajo:

- Conocer qué es el mundo Big Data y Business Intelligence.
- Conocer de primera mano cómo trabaja un club de fútbol profesional en el análisis de los datos.
- Conocer las principales plataformas de obtención de estos datos.
- Analizar una gran cantidad de datos a partir de diferentes métricas y KPI's.
- Crear un algoritmo de similitud que permita comparar el rendimiento de un jugador con el resto.
- Conocer el programa Power Bi en su conjunto, así como aprender a realizar un buen uso de este.
- Realizar un supuesto práctico dentro del mundo real, para así poder conocer de primera mano los puntos mencionados anteriormente.

1.4 Descripción del trabajo planificado para alcanzar los objetivos propuestos

1. Investigación y estudio sobre Big Data y Business Intelligence:

Leer libros, artículos y recursos en línea para comprender los conceptos fundamentales y las aplicaciones del Big Data y Business Intelligence en diferentes industrias.

Participar en cursos en línea o presenciales sobre el tema para adquirir conocimientos más profundos y actualizados.

2. Estudio del análisis de datos en un club de fútbol profesional:

Realizar investigaciones y leer estudios de casos sobre cómo los clubes de fútbol profesional utilizan el análisis de datos para mejorar su rendimiento.

Intentar establecer contactos con profesionales del análisis de datos en el ámbito futbolístico para obtener información de primera mano y posibles oportunidades de aprendizaje.



3. Investigación de las principales plataformas de obtención de datos:

Identificar y explorar las plataformas más utilizadas en la obtención de datos en el ámbito del fútbol y el análisis deportivo.

Analizar las características, funcionalidades y ventajas de cada plataforma para determinar cuál se adapta mejor a tus necesidades y objetivos.

4. Análisis de datos utilizando diferentes métricas y KPI's:

Identificar las métricas y los indicadores clave de rendimiento (KPI's) relevantes para el análisis de datos en el fútbol.

Recopilar y analizar una gran cantidad de datos utilizando herramientas de análisis y visualización de datos específicas para el deporte.

5. Desarrollo de un algoritmo de similitud:

Estudiar los algoritmos existentes utilizados para comparar el rendimiento de los jugadores en diferentes deportes.

Diseñar y desarrollar un algoritmo de similitud que permita comparar el rendimiento de un jugador con respecto a otros, considerando las variables relevantes.

6. Aprendizaje y práctica de Power BI:

Realizar tutoriales y cursos especializados para familiarizarse con Power BI y aprender a utilizar todas sus funciones y capacidades.

Crear visualizaciones y paneles interactivos para presentar los datos y el análisis de manera efectiva.

7. Realización de un supuesto práctico en el mundo real:

Buscar oportunidades para colaborar con equipos deportivos, instituciones o empresas relacionadas con el análisis de datos en el fútbol.

Aplicar los conocimientos adquiridos para realizar un proyecto práctico que aborde los puntos mencionados anteriormente, utilizando datos reales cuando sea posible.



1.5 Motivación

Hay personas que nacen con un propósito de vida definido, yo nací con la camiseta del Numancia. El fútbol me ha acompañado desde bien pequeño, debido a que mi abuelo, socio del Numancia desde los 12 años, así me lo inculcó.

Durante toda mi juventud me dediqué a jugar al fútbol, desde el equipo de mi colegio hasta jugar en las categorías inferiores del club de mi ciudad. Esto, junto que mis padres tienen varios amigos que viven del fútbol como profesión, desataron en mi esta pasión. Con el objetivo de poder en un futuro dedicarme profesionalmente a ello, decidí formarme como entrenador, obteniendo el título a los 16. Debido a que, en este mundo, si quieres destacar, debes de ser diferente, me apunté a un curso de Big Data deportivo hace unos meses, lo cual desató en mi un camino el cual seguir para conseguir mi sueño.

Debido a esta pasión y los consiguientes conocimientos que he podido adquirir al cursar tanto el grado en Organización Industrial, como el curso en Big Data mencionado anteriormente, decidí realizar este trabajo de fin de grado. El objetivo es poder seguir divulgando en el mundo de los datos y así, seguir obteniendo una formación autodidacta, además de poder demostrar que el Big Data y la inteligencia de negocio son el futuro en el mundo del fútbol.



2. BIG DATA

2.1 ¿Qué es el Big Data?

Big data (grandes datos), como su propio nombre indica, consiste en grandes volúmenes de datos, da igual en el ámbito que se aplique. Estas cantidades de datos, cada día que pasa son mayores, ya que va avanzando la tecnología y con ello la capacidad de poder medirlos.

Buscando en internet definiciones más concretas, podemos añadir las siguientes:

"Un gran volumen, velocidad o variedad de información que demanda formas costeables e innovadoras de procesamiento de información que permitan ideas extendidas, toma de decisiones y automatización del proceso"

(Gartner- consultoría tecnológica, 2020)

"Al análisis de información en cantidades industriales, todo un universo de oportunidades para las empresas aún por explorar"

(Empresas Forbes, 2021)

2.2 Historia del Big Data

El inicio del Big Data no tiene una fecha concreta de creación, debido a la cantidad de datos a nivel global que se han podido recoger y tratar en cada momento de la historia no ha sido el mismo, y a que es difícil señalar, a partir de qué punto, podemos considerarlos realmente como "una gran cantidad de datos".

Es por ello, bastante complicado poner una fecha exacta a la creación de este término, pero existen autores que sitúan la primera aparición del Big Data en la década de los años 90. Esta fecha coincide con el inicio de Internet. Otras personas dicen que estos términos de grandes datos provienen desde épocas del paleolítico, debido a que los seres humanos siempre hemos sido una especie con un evidente interés en la recopilación de información y extraer conocimiento de ella.

Ha sido el desarrollo de diferentes aplicaciones informáticas para la recogida de esos datos y, los avances continuos en la tecnología de análisis de datos realizados en los últimos años, lo que ha permitido dar forma a las prácticas que ahora conocemos como Big Data.

2.3 Nacimiento del término Big Data

La palabra "Big Data" la utiliza por primera vez Erik Larson (periodista estadounidense) en el año 1989 en un artículo que escribe sobre el marketing y como se podrán usar los datos de diferentes clientes.



Con la llegada de internet en los años 90, como hemos mencionado anteriormente, que permitía interconectar en red múltiples computadoras, el término Big Data empieza a tener cierto auge. Este se vio amplificado debido a que con la llegada del primer navegador web para la WWW (World Wide Web), de forma que cualquier usuario en cualquier lugar del mundo, con un acceso a la red, podía subir y hacer disponibles los datos que quedaban almacenados y compartidos en los servidores. Esto supuso que se empezase a generar y tener disponible una gran cantidad de datos.

Es en la década de los 2000, junto con la creación años antes del buscador Google, cuando surgen una gran cantidad de empresas dedicadas a generar, tratar y almacenar de forma segura grandes cantidades de datos. Además, en 2005 nace la denominada web 2.0, donde se facilita la forma de compartir información, y, por tanto, aumenta la más cantidad de datos generados.

En los años siguientes, y más cercanos a la actualidad, donde se produce el florecimiento de la disciplina del análisis de grandes cantidades de datos, haciendo que el número de personas y empresas que se encargan exclusivamente de la recopilación y gestión de los datos crezca de manera exponencial. Se genera así una gran cantidad de empleos y surgen empresas exclusivamente centralizadas en la obtención de datos, prácticamente, de cualquier tipo.

2.4 Big data en la actualidad. (Beneficios y usos)

Actualmente, con tan solo una simple búsqueda en internet, podemos encontrar una gran cantidad de datos disponibles sobre cualquier tema, sector o disciplina. En la actualidad, la recopilación de los datos que generan ellas o en su entorno, y su tratamiento posterior, son un recurso muy valioso para la mayoría de las empresas. Su principal objetivo es el análisis y procesado de todas estas grandes cantidades de datos hasta convertirlos en información que les resulte útil como apoyo para la toma de sus decisiones.

“Sin Datos eres sólo otra persona más con una opinión”

(William Edwards Deming (1900-1993))

Algunos de los principales beneficios que podemos encontrar en el Big Data son los siguientes:

1. Minimizar el error humano y encontrar oportunidades de mejora en diversas áreas.

Se miden diferentes datos de situaciones similares que se han realizado en el pasado. Con estos datos se analizan todos los departamentos y se estudia las áreas en los que se puede inculcar una mejoría. Se pasa de una simple intuición a un proceso analítico. Con ello se logra minimizar el error humano.



Ejemplo: Se ve que en un proceso de fabricación hay una parte del proceso que registra una variabilidad en los tiempos de ejecución muy por encima de las demás. Se analizan esos datos, en este caso, buscando el posible origen de esa variabilidad en los tiempos, y se proponen soluciones para minimizar el error y así mejorar la efectividad en la producción.

2. Aumentar la velocidad de la toma de decisiones

A través del Big Data se pueden estudiar la existencia de diferentes patrones repetitivos en determinados comportamientos. Por ejemplo, podemos ver que busca cada cliente en su celular o, mismamente, cómo reacciona cada uno de ellos a una campaña de marketing. También podemos medir datos sobre el uso que realizan de los diferentes productos comprados en una empresa y, así, analizar cuáles son los puntos fuertes o sus puntos débiles de esos productos.

Con todos estos datos recopilados, y un adecuado procesado posterior de los mismos, se podría, por ejemplo, poner a disposición de los directivos una herramienta, como un buen cuadro de mando para la visualización de resultados obtenidos, que agilice la interpretación de esos datos y que así puedan realizar una toma de decisiones lo más rápida posible.

3. Conocimiento más amplio de los clientes

Las empresas con un sistema de Big Data analítico implementado pueden llegar incluso a saber el porqué de las compras que realiza cada cliente. Siguiendo la línea del apartado anterior, conociendo los gustos de los clientes y las variables relacionadas con sus compras, se puede analizar cómo o cuándo van a adquirir los productos o servicios.

Ejemplo: Un cliente se compra una camiseta blanca. Este cliente al cabo del tiempo se pregunta porque se compró una camiseta blanca y no una negra. Y no obtiene la respuesta. Sin embargo, la empresa que ha vendido esa camiseta, analizando los datos recopilados y su correlación con otros como pueden ser su personalidad, últimas compras, época del año, o simplemente lo que el cliente comenta en redes sociales, si puede llegar a saber esas razones con cierta fiabilidad.

Y conociendo mejor a los clientes, se puede planificar el inventario requerido, las adquisiciones de productos o materia prima a proveedores, ... y así crear una estrategia mejor enfocada a la hora de satisfacer las preferencias en el mercado en cada momento.

4. Optimización de procesos y reducción de costos

A partir de grandes cantidades de datos previos, por ejemplo, financieros, se pueden simular escenarios para combinarlos, analizando el what-if (que pasaría si), con análisis de los datos basados en el machine learning con la finalidad de comprobar diferentes variables de negocio. Con ello se optimizan los procesos y se reducen los costos de una manera rápida y precisa.



2.5 Big Data en el futuro

El año pasado (2022) ha sido un año de asentamiento y consolidación del personal profesional en Big Data dentro de las empresas, ya que estas, están demandando notoriamente perfiles de científicos de datos y analistas de datos. Con estos perfiles buscan como objetivo obtener un máximo rendimiento de los datos para así, poder generar un impacto positivo dentro de sus diferentes organizaciones.

“Los datos son ‘el nuevo petróleo’, se están convirtiendo en pieza clave de la sociedad y la economía”

(Antonio Conde (director de Cisco España), 2022)

Estos datos, son los nuevos activos por explotar dentro de la empresa, y las que consigan dominar su tratamiento y análisis, tendrán una ventaja respecto al resto. Por ello, la mayoría de las empresas van a implementar los siguientes puntos en el 2023.

1. Data Strategy

Hasta ahora la estrategia de datos de las empresas se formulaba dentro de los departamentos de tecnologías de la información o bien de departamentos propios como “Data”. Para este año, la tendencia parece haber cambiado, ya que la mayor parte de las empresas introducen sus proyectos el análisis de datos como parte de sus estrategias primarias, en lo que se denomina como “core bussiness” (actividades de gran valor o principales dentro de la organización), como Data Strategy.

Por tanto, en 2023 y a lo largo de años posteriores podremos comprobar como todas las empresas van a incluir en sus planes estratégicos en mayor medida elementos de Data Strategy.

2. Inteligencia Artificial

“La inteligencia artificial (IA) es el campo de la ciencia informática dedicado a la resolución de problemas cognitivos asociados comúnmente a la inteligencia humana, como el aprendizaje, la resolución de problemas y el reconocimiento de patrones. La inteligencia artificial, que normalmente se abrevia IA, puede evocar escenas futuristas o robóticas, pero va mucho más allá de los modelos autómatas de ciencia ficción, ya que se trata de ciencia informática avanzada actual.”

(Amazon IA, 2021)

En otras palabras, y resumidamente, la Inteligencia Artificial se encarga de aprender de los errores producidos en cualquier ámbito y simular la inteligencia humana en máquinas, dando la capacidad de aprender, razonar, hacer predicciones... y lo hacen

en base a los datos disponibles. Cuanto mejores sean los datos y la cantidad de estos, mejores son los resultados.

La inteligencia artificial está permitiendo que haya una mejora en la velocidad de la toma de decisiones y, por tanto, las empresas que adopten e incorporen esta inteligencia artificial dentro de sus procesos de negocio, obtendrán una ventaja competitiva respecto al resto de empresas. De hecho, ya hay empresas que gestionan ciertos ámbitos de estas exclusivamente con inteligencia artificial.

En la Imagen 1, podemos comprobar los distintos usos cotidianos de la inteligencia artificial.



Imagen 1. Distintas aplicaciones de la IA. Fuente: (Europa.rl, 2021)

3. Datos y almacenamiento en la nube

Una de las tendencias actuales del Big Data es operar dentro de los servicios de la nube o iCloud. Esto hace que cualquier usuario o aplicación pueda tener acceso a un montón de información de forma distribuida, sin depender de donde puedan estar estos usuarios o aplicaciones.

4. Democratización de los datos

Otro cambio que se espera dar en las empresas en este 2023, es que todos los departamentos de la empresa puedan trabajar con datos y no solo el departamento de Data Scientist o el departamento de Big Data.



Un estudio realizado por McKinsey (consultora estratégica global) dio a conocer que, si en una empresa todos los interesados tienen acceso a los datos y trabajan adecuadamente con ellos, existe una probabilidad 40 veces mayor de tomar las decisiones acertadas.

5. Big data y cambio climático

Puede que el cambio climático no sea un tema nuevo para los científicos, pero el Big Data se posiciona en el 2023 como la principal herramienta para ayudar a conocer el estado actual de las emisiones de Dióxido de carbono.

“En este contexto de alteraciones, la capacidad de recopilar e interpretar la enorme cantidad de datos sobre el clima a largo plazo y de forma continua representa una herramienta fundamental para entender si los fenómenos meteorológicos excepcionales como los referidos pueden considerarse eventos aislados o ser indicadores de importantes cambios en las condiciones climatológicas del planeta”

(Carlo Buontempo – director de C3S, 2023)

C3S (Servicio de cambio climático de Copernicus) es un organismo que se encarga de poner a disposición de la sociedad, información sobre la situación climática pasada, presente y de futuro del planeta. Además, se encarga de procesarla con técnicas de Big Data, para convertir esta información masiva en indicadores fáciles de entender y ofrecerlos de manera gratuita a diferentes países de la Unión Europea. Es un referente muy importante para nuestro país, debido a que España es uno de los países europeos con mayor diversidad climática.



3. INTELIGENCIA DE NEGOCIO (BUSINESS INTELLIGENCE)

3.1 ¿Qué es el Business Intelligence?

Business Intelligence (BI) o inteligencia de negocio, es el conjunto de procesos necesarios para transformar los datos generados en uno o varias áreas de negocio en conocimiento, de forma que este permita una mejor toma de decisiones dentro de las empresas.

"Aunque la BI no les dice a los usuarios empresariales qué hacer o qué sucederá si toman un determinado curso, la BI tampoco se trata únicamente de generar informes. Más bien, BI ofrece una forma para que las personas examinen los datos para comprender las tendencias y canalizar los insights"

(Revista CIO, 2022)

La inteligencia de negocio surgió en el 1960 como un sistema para compartir la información entre diferentes empresas. Más adelante, en la década de los 80 se desarrolló mucho más debido a las primeras apariciones de modelos informáticos. Durante esos años, únicamente la finalidad de la inteligencia de negocio era transformar los pocos datos que se tenían en información para la ayuda de toma de decisiones.

Actualmente la inteligencia de negocio prioriza además otros factores como el análisis de autoservicio flexible, los datos gobernados en plataformas confiables, la capacitación de los usuarios corporativos y la rapidez para obtener información.

3.2 Funcionamiento de la inteligencia de negocio

Las diferentes plataformas de inteligencia de negocio se basan en la existencia de unos almacenes de datos para recoger la información básica, y generalmente masiva, obtenida desde diversos orígenes de datos. Desde este almacén de datos, el sistema se encarga de suministrar, filtrar, agregar y procesar si es necesario, esos datos hasta una interfaz de usuario que, a través de diferentes visualizaciones, ya sea en forma de tablas, gráficas, animaciones... y con un mayor o menor grado de interactividad, permiten profundizar en diferentes aspectos sin tener que procesar los datos a bajo nivel. Cabe destacar que un inconveniente de las herramientas de Business Intelligence actuales es que aún presentan limitaciones notables a la hora de trabajar con la gran cantidad de datos que se requiere procesar en un sistema actual de Big Data.

Es por lo que, a la hora de implementar los almacenes de datos, se recurre a incluir un motor OLAP (motor de procesamiento analítico en línea), que admita cualquier tipo de consulta. Con él se procesan ciertas consultas preestablecidas de antemano (filtrados, segmentaciones de datos, etc.), cuyos resultados agregados son los datos que realmente son almacenados. Se consigue así, aligerar la carga de datos que almacenará el sistema, al no guardar todos los datos masivos a bajo nivel, sino sólo los valores agregados.



Además, se consigue acelerar el funcionamiento general al no tener que volver a realizar esas consultas. La contrapartida es la pérdida del nivel de detalle en los datos.

"OLAP proporciona una poderosa tecnología para el descubrimiento de datos, lo que facilita la inteligencia empresarial, los cálculos analíticos complejos y el análisis predictivo"

(Doug Dailey (gerente de ofertas de IBM – empresa tecnológica), 2021)

Esto lo entenderemos mejor a través del siguiente ejemplo:

¿Cuáles fueron las ventas en cada región española este año, respecto al año anterior?

Es una posible pregunta que se pueden realizar todas las empresas que operen en España. El entorno de Business Intelligence se encargará de ir al sistema de almacén de datos origen y recoger todas y cada una de las ventas tanto del año pasado, como del presente año que está finalizando. Podrían ser millones de registros. Esos datos deben ser procesados en ese momento en función de la consulta realizada por el usuario, y con los resultados de la consulta, el sistema genera una serie de informes visuales donde se podrían comparar diferentes tipos de datos que podrían ser de gran interés, como cifras de ventas totales, cifras de ventas por regiones, fecha de las ventas...

Pero si las consultas se basan siempre en unos mismos filtros y segmentaciones, entonces los datos que podríamos guardar podrían ser simplemente los resultados de las combinaciones posibles de esos filtros y/o segmentaciones, no el total de esos millones de datos, con el consiguiente ahorro en tiempo y espacio de almacenamiento.

Esto acelera la creación de los diferentes informes visuales personalizados que mandaríamos a los diferentes jefes de los departamentos, los cuales son los encargados de tomar decisiones. En función de los resultados obtenidos en estos informes generados por la inteligencia de negocio, los responsables de la empresa deberían ser capaces de tomar las mejores decisiones posibles para la misma.

Algunas soluciones de inteligencia empresarial más actuales pueden extraer e ingerir datos sin procesar directamente, utilizando otro tipo de tecnologías como Hadoop (software que permite programar aplicaciones que soporten grandes volúmenes de datos), pero los almacenes de datos siguen siendo la fuente de datos de elección en muchos casos.

3.3 Ventajas de la Inteligencia de negocio

En la Imagen 2 se pueden observar las principales ventajas de utilizar la inteligencia de negocio en la toma de decisiones.

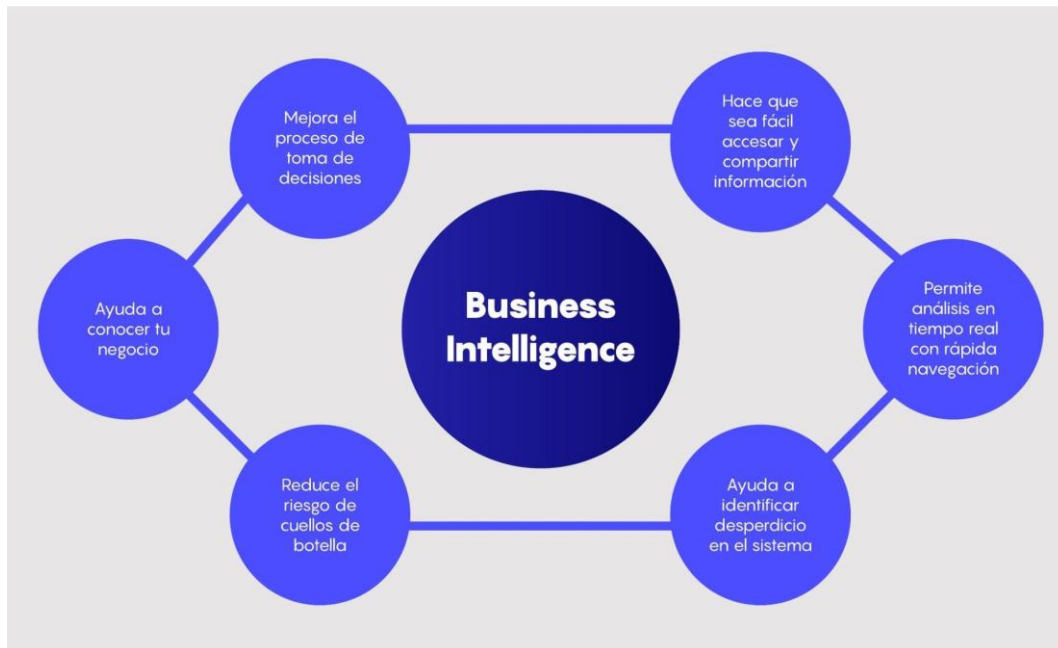


Imagen 2. Ventajas del Business Intelligence. Fuente: (Blog.Interius)

1. Cuadros de mando interactivos

Una de las herramientas claves del Business Intelligence es la elaboración de cuadros de mando interactivos, que son fáciles de usar y que son capaces de recoger datos en tiempo real desde diferentes fuentes de almacenes de datos. Estos cuadros de mando interactivos suelen permitir clicar en las propias representaciones, de forma que estas funcionan como filtros o segmentaciones de datos en función de lo seleccionado. Y no sólo para la propia visualización, sino también para el resto de las presentes en el cuadro vinculadas a ella. De esta forma, todo se actualiza en función del filtro o segmentación seleccionado, haciendo que aparezcan/desaparezcan de las tablas, gráficas... los datos no seleccionados, o que se estos se agreguen / desagreguen... Mediante esto, el usuario puede indagar en los datos al mismo tiempo que observa los cambios en las visualizaciones. Esto lo podemos ver con la siguiente Imagen 3.

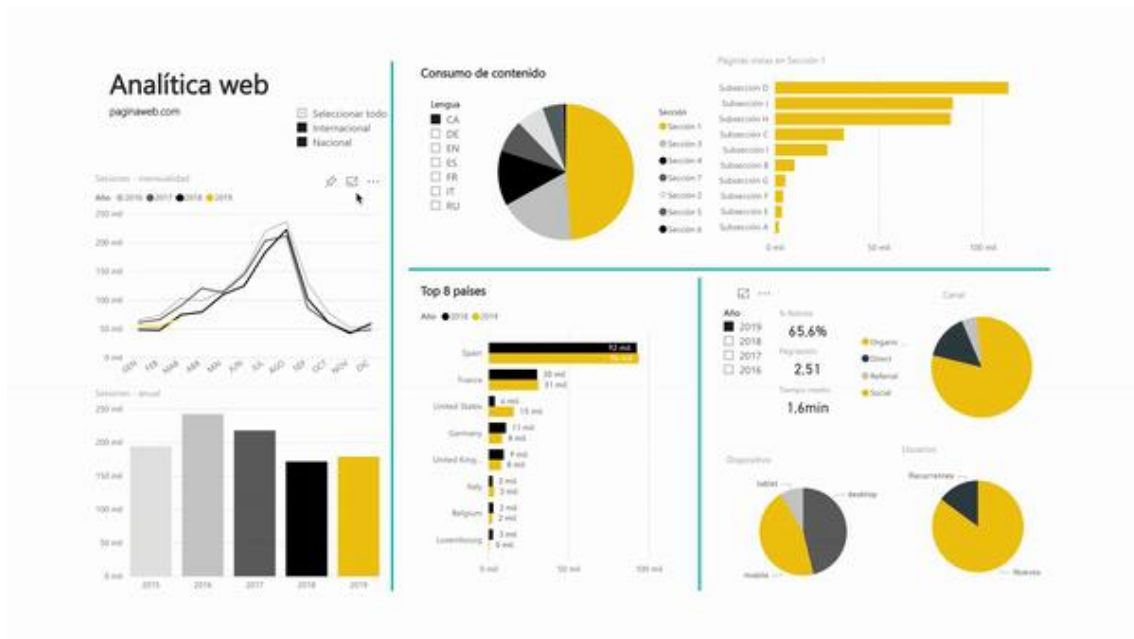


Imagen 3. Ejemplo de cuadro de mandos interactivos. Fuente: (La magnética)

2. Aumento de la eficiencia

La mejora de la eficiencia (capacidad para realizar o cumplir adecuadamente una función) es uno de los principales beneficios del BI dentro de la empresa. Esta gestión de los datos permite obtener a la alta dirección de la empresa una visión más global, con la que poder identificar los principales problemas de esta.

Este aumento de la eficiencia, liberándolos de tareas de procesamiento recurrente de datos, permite a los empleados disponer de un mayor tiempo para que se centren en otros trabajos de innovación y creatividad.

3. Proceso inteligente de toma de decisiones

Como ya hemos mencionado en el primer punto, una buena implementación de BI proporciona datos precisos que se pueden actualizar/analizar incluso en tiempo real. Esto permite a la mayoría de las empresas una toma de decisiones inteligente enfocada en la predicción, el presupuesto y la planificación con los datos más actualizados posibles.

Así, una empresa podría, por ejemplo, consultar las previsiones actualizadas de las necesidades y deseos de los clientes, antes incluso de reunirse con ellos.

4. Satisfacción del cliente

Una combinación acertada de la inteligencia artificial y de las herramientas de inteligencia de negocio en una empresa ayudan a ajustar mejor las satisfacciones de los clientes en cuanto a gustos y también a precio. Esto hace que los clientes sean fieles a la empresa y así, poder conseguir mayores ingresos en un futuro.



5. Extracción de datos con mayor precisión

La “minería de datos es una tecnología que nos permite trabajar con grandes cantidades de bases de datos de manera automática o semiautomática, con el objetivo de encontrar patrones, tendencias o reglas que nos ayuden a comprender el comportamiento de los datos.”

(Rivera Resina, 2018)

El denominado Data Mining o minería de datos busca obtener patrones, tendencias y reglas en grandes cantidades de datos, y sigue cinco pasos: la recopilación, el almacenamiento, la organización, el análisis y la visualización de datos. Todo el proceso se debe poder completar sin ningún error potencial y sin intervención humana, para evitar graves consecuencias como una pérdida de datos en el futuro.

Se trata pues de un proceso análogo en sus fundamentos a la Inteligencia de Negocio, de forma que puede ser previo /o implementado conjuntamente con del sistema de Inteligencia de Negocio, de forma que después podamos aprovechar esos patrones y tendencias detectados, como una parte más en la toma de decisiones propia del sistema de Inteligencia de Negocio.





4. BIG DATA DEPORTIVO

4.1 ¿Qué es el Sports Analytics?

El Sport Data Analytics (análisis de datos en el deporte) es la disciplina que utiliza grandes volúmenes de datos para lograr diferentes objetivos.

Cada partido en la actualidad puede generar hasta más de 8 millones de datos. Lo que se pretende es encontrar patrones repetitivos tanto del propio equipo, como del equipo rival para ayudar a la toma de decisiones. La clave es tener cuántos más datos posibles mejor, para preparar cualquier situación que se nos pueda dar durante el partido o competición, sin antes olvidar, que el deporte, por mucho que se analice, depende de los deportistas y no de los números.

"Tienes que saber que dato te interesa. En cada partido hay miles de datos. Esos datos se pueden tratar para sacar conclusiones. No todo te sirve. Si quiero saber el estado físico de mi equipo, me interesa saber cuánto he hecho en velocidad o en volumen"

(Ramón Rodríguez (Monchi), 2020)

Cada vez más clubes y más deportes recurren a este tipo de análisis en sus departamentos deportivos.

4.2 Orígenes del Big Data en el deporte

4.2.1 Sabermetría

La sabermetría no es más que el análisis de los datos en el béisbol a través de las estadísticas. Se puede considerar como los inicios del Big Data en el deporte, alrededor de los años 70. Bill James, un periodista de la época, empezó a realizar y analizar los registros históricos de los jugadores de Beisbol de las grandes ligas. Sus artículos trataban sobre los datos estadísticos de cada jugador y no sobre criterios puramente subjetivos como se realizaba de manera habitual.

En 1985 publicó un libro denominado 'Baseball abstract' donde recogía todos esos datos y artículos.

4.2.2 Moneyball

Es una adaptación de la película 'The art of winning an unfair game en 2003' que se basa en la historia real de Billy Beane, gerente general del equipo Oakland Athletics, quien utilizaba las estadísticas avanzadas para fichar jugadores.

La película (Imagen 4) nos cuenta como Billy Beane pasó de configurar la plantilla del Oakland Athletics en función del scouting (búsqueda) a fichar jugadores basándose en

los fundamentos de la sabermetría explicados en el punto anterior. Billy lo que pretendía demostrar era como era igual o más importante ver y analizar los datos de los jugadores, antes que ficharlos únicamente por lo que podíamos apreciar a través del ojo humano. Esto le llevo a enfrentarse con su plantilla de ojeadores. Su objetivo era poder construir una plantilla competitiva en función de la calidad/precio de cada uno de sus jugadores.

Tras este cambio en la manera de realizar los fichajes para la entidad, los Oakland Athletics fueron los primeros en ganar 20 partidos seguidos en toda la historia del Beisbol.

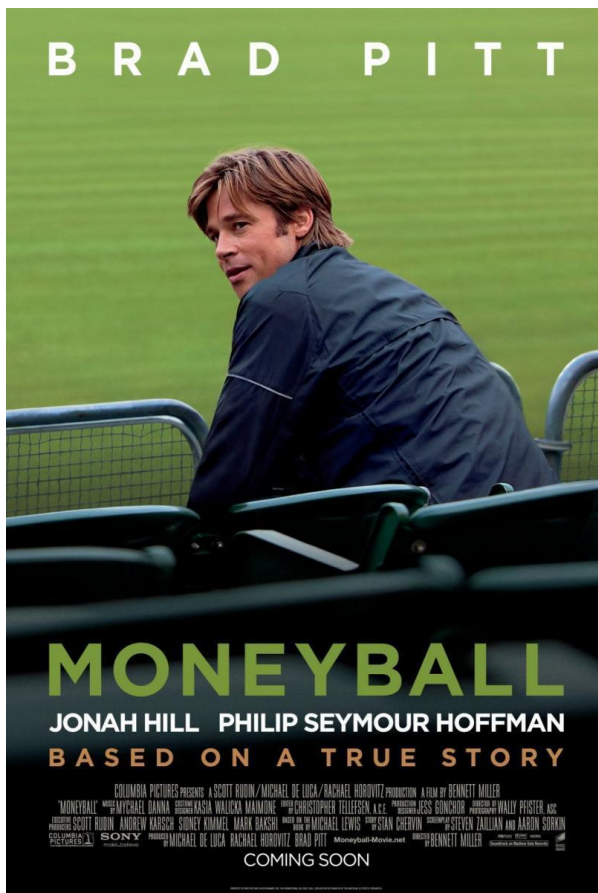


Imagen 4. Película Moneyball (2011).
Fuente: (FilmAffinity)

4.2.3 Muthuball

Muthu Alagappan se graduó en la universidad de Stanford en 2012 como ingeniero biomédico y en 2016 se graduó como médico. Durante su etapa como alumno, diseño un software, en el cual presentó 13 posiciones contra las 5 clásicas que tiene un equipo de baloncesto. Empezó a redefinir a través de los datos, el modo en el cual se posicionan y se construyen los equipos. Su principal base, es que todo gira en torno al jugador y su forma de jugar, y para ello busca las diferentes posiciones que describan de la mejor manera posible la forma de jugar.



Este estudio y presentación le sirvió más adelante para fichar por equipos como los Portland o Miami Heat. Actualmente, este estudio sirve para la planificación y construcción de plantillas.

4.3 La importancia del Big data en el deporte

El Big data en el deporte como se ha dicho a lo largo del presente trabajo, es fundamental en los tiempos que corren hoy en día, ya que es una herramienta fundamental para el análisis en cualquier ámbito y no lo va a ser menos en el del deporte.

El primer paso por identificar es como se utiliza la estadística dentro del deporte y cuál es la importancia que se le debe de dar.

“Lo que se hace en estadística es identificar variables de una población, en el deporte serán las jugadas, y si pensamos en una ciudad pueden ser personas, y de ahí sacamos unos valores que son nuestra muestra. Entonces lo que hace la estadística es identificar una muestra con pocos valores, es decir, como los datos que manejamos son tan grandes, vamos a tomar esta muestra y la hacemos igual a un valor para que nos empiece a decir cosas. Si nos planteamos un interrogante inicial y con los datos intentamos responderla, luego seguramente aparecerán otras. La pregunta es ¿Qué hace bueno a jugadores a Neymar, Luka Modric, Lebron James o Tom Brady?”

(Jesús Soler (Ingeniero Español), 2020)

Esta pregunta es muy difícil de responder, ya que cada uno tendríamos una opinión muy subjetiva, pero... analizando y estudiando todas estas estadísticas (datos) podríamos llegar a dar una respuesta objetiva a la pregunta.

Actualmente, en el fútbol se están probando nuevas métricas de análisis, como podemos comprobar en cualquier partido visto por la televisión. Estas, se centran en analizar el disparo y calcular cual es la probabilidad de anotar gol en función de metros de distancia a portería, ángulo, otros disparos similares hechos por ese jugador en esa posición o número de rivales entre la zona de disparo y portería.

Por ello, con esta métrica se le puede avisar al jugador de cuál es la posición más probable para poder anotar gol y así poder ayudar a ganar un partido de fútbol.

En el fútbol el Big Data está cogiendo cada vez más relevancia, pero aún no se llega a masificar a un nivel extremo.

En la NBA (baloncesto) ocurre todo lo contrario. Todos los equipos tienen un sistema de Big Data y un área de recolección de datos de los diferentes jugadores de la competición. Esto ocurre desde 2007 cuando el Ingeniero Daryl Morey decidió cambiar la mentalidad del baloncesto. Su plan consistía en: Jugar rápido para intentar tirar en menos de 10 segundos, tener más lanzamientos triples que dobles, Pick & Roll (bloqueo al jugador

exterior para conseguir ventaja) y posicionamiento de jugadores abiertos más allá de la línea de 3 puntos para recibir el pase. Los Rockets, que era el equipo para el cual trabajaba Daryl Morey, decidió fichar a jugadores específicos para este tipo de juego, analizando los datos de las métricas comentadas anteriormente. Con ello, los Rockets pasaron a ser uno de los equipos más importantes de la competición.

En la actualidad, la NBA maneja aplicaciones Big Data. Una de ellas es el “Ghosting” la cual se encarga a través del análisis de datos de realizar un software con el cual se produzcan simulaciones de situaciones reales que se pueden dar dentro de un partido de baloncesto, para así poder entrenar este tipo de situaciones y que a los jugadores les parezcan más cercanas cuándo se den dentro de un partido en la realidad.

Otra aplicación que maneja la NBA es la Homecourt (Imagen 5). Esta se encarga de darnos datos en tiempo real y de analizar y predecir posibles fichajes dependiendo de las virtudes de cada jugador y de la probabilidad de que se adapte bien al equipo.

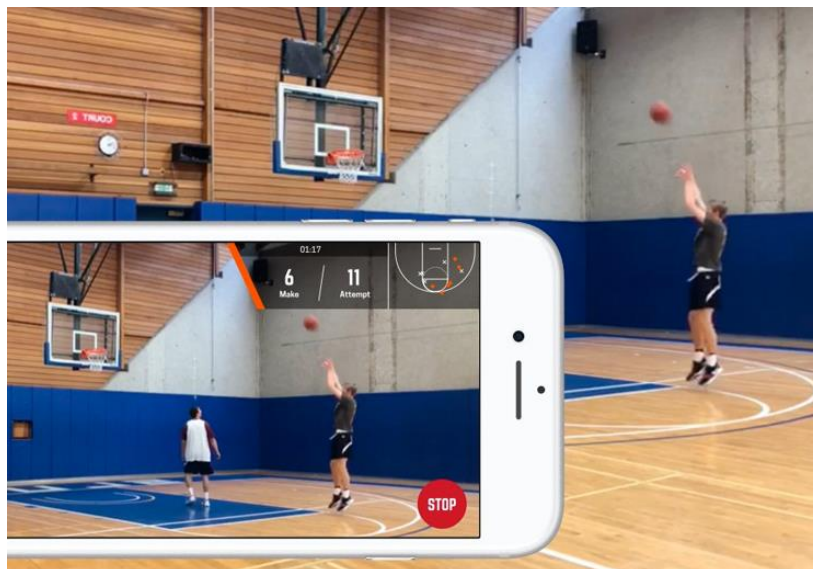


Imagen 5. Aplicación HomeCourt. Fuente: (VivaBasket)

Otro de los deportes en el cual se utiliza el Big Data de manera masiva es el Rugby. Los analistas y aplicaciones como Opta hacen que se pueda observar las diferentes zonas del juego, identificar jugadores clave del oponente...

Además, mediante dispositivos GPS se permite que se puedan sacar datos físicos específicos de cada jugador, para así poder ayudar en la reducción de probabilidad de lesiones.

La selección de Nueva Zelanda, más conocida como los “All Blacks”, utiliza el mismo método mediante “SAS Visual Analytics” que se encarga de recolectar todos los datos de jugadores que juegan tanto en su país como en el extranjero con la finalidad de convocar a los mejores para representarlos en las diferentes competiciones.

"No es lo mismo que no sepas qué va a hacer el rival a que tengas una idea muy clara de lo que hace en diferentes situaciones. No me gusta decir que la estadística gana partidos, pero es una herramienta esencial para que vos entiendas qué es lo que pasa y lo que va a pasar".

(Genaro López (Analista de datos Federación vóley Argentina), 2019)

El Big Data se posiciona como una herramienta fundamental en aportar gran información a los diferentes clubes y selecciones para poder sacar ventajas sobre los competidores. En mi opinión y como he puesto en la primera línea del presente párrafo, es sólo una herramienta, ya que el deporte es mucho más que eso y depende de miles y miles de factores. Si únicamente dependiese de los datos, no se llamaría deporte.

"Íbamos a trabajar y hacíamos un listado de laterales, le dábamos a Monchi el listado. Pasaban dos semanas en las que él analizaba el listado. Venía y decía: el primero es el mejor lateral. Pero el tercero es el mejor para el Sevilla. Yo me envenenaba, porque luego, cuando empezaba a discutirlo y hablar con él, tampoco me daba una razón. Y acertaba siempre"

(Víctor Orta (director deportivo Leeds), 2021)



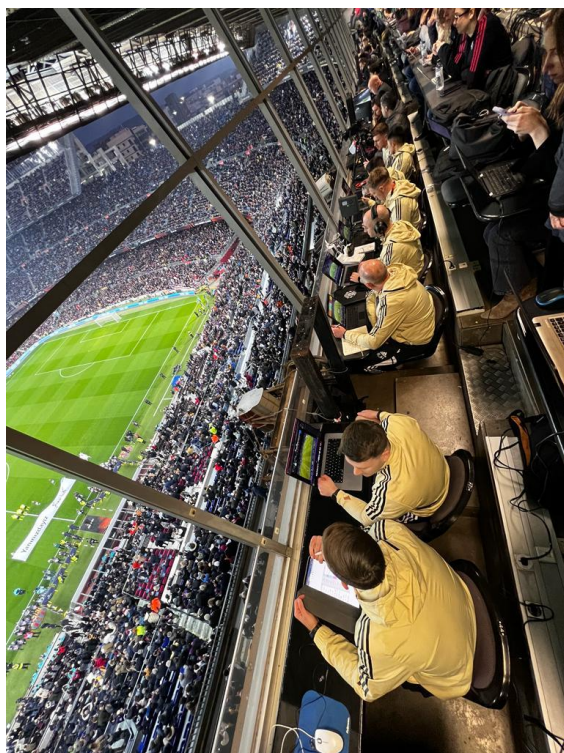
Imagen 6. Charlando con Víctor Orta en M2 Máster scouting. Fuente: (Elaboración propia)



5. BIG DATA EN EL FÚTBOL

5.1 Usos del Sport Analytics en el mundo del fútbol

En la actualidad el uso del Big Data dentro del mundo del fútbol ya es una realidad. En los clubes de ligas de alto nivel ya hay grupos de personas especialistas que se dedican exclusivamente al análisis de estos datos. Esto lo podemos ver de una manera muy representativa en la siguiente Imagen 7, correspondiente al partido de ida de dieciseisavos de Europa League entre el FC Barcelona y Manchester United.



*Imagen 7. Analistas Manchester United
Partido Europa League. Fuente:
(VarskySports – Twitter)*

Se analizan multitud de parámetros, los cuales podemos en los siguientes campos:

1. Análisis técnico – tácticos

Se analizan datos correspondientes a acciones que suceden dentro del terreno de juego. Pases, tiros, regates con éxito, posesión de balón, toque de cada jugador en campo rival...

2. Análisis físicos

Se analizan datos físicos como pueden ser, la distancia recorrida por cada jugador, la velocidad máxima, número de sprints realizados...



3. Búsqueda de nuevos jugadores:

A través de una serie de parámetros y diferentes KPI creados por el correspondiente grupo de trabajo de la dirección deportiva.

4. Prevención de lesiones:

Se utilizan los datos físicos para crear nuestros propios KPI para estudiar las posibles lesiones.

Además de los campos mencionados anteriormente, el Big Data dentro del fútbol también se utiliza para diferentes factores de fuera del terreno de juego. Se utilizan como si fuese una empresa de cualquier sector, para analizar el número de entradas vendidas, informes económicos dentro del club, redes sociales o incluso informes de recursos humanos.

5.2 Proveedores de datos y microchips GPS

Al ver un partido de fútbol se puede observar que los jugadores, debajo de las camisetas, llevan un chaleco en el cual se recoge un chip. Este chip es el denominado PLAYR y es un sistema SmartCoach. Ver Imagen 8.

El sistema tiene tres componentes: SmartPod (el dispositivo en la espalda), SmartVest (el chaleco) y la aplicación SmartCoach. Los tres juntos presentan información relevante del rendimiento de cada jugador en las sesiones de entrenamiento o partidos, midiendo velocidad, distancias recorridas, creando además mapas de calor.

Con este sistema se pueden medir todos los datos físicos que registra cualquier jugador en cada entrenamiento o partido.



Imagen 8. PLAYR (SmartCoach). Fuente: (todosobrecamisetas)

Todos estos datos registrados por este SmartCoach, además de todos los datos técnico – tácticos de cada partido se recogen de manera pública (páginas web) como puede ser transfermarkt o fbref, o mediante a una suscripción a empresas proveedoras de datos, como son Instat, wyscout o mediacoach. Este último es el proveedor oficial de datos de los equipos de Primera y Segunda división.

Si tenemos la opción de tener varios proveedores de datos como podemos ver en Imagen 9 e Imagen 10, más podremos contrastar nuestra veracidad y la calidad del dato será próxima al 100% de autenticidad.



Imagen 9. Principales proveedores de datos técnico-tácticos. Fuente: (ObjetivoAnalista)



Imagen 10. Principales empresas que suministran dispositivos SmartCoach. Fuente: (ObjetivoAnalista)

5.3 Eventing y Tracking

Los datos de Eventing y Tracking son las principales formas mediante las cuales se obtienen los datos en el mundo del fútbol.

“El uso de datos de Eventing y Tracking, además de permitirnos ahorrar tiempo en el análisis, nos permite reducir la posibilidad de extraer conclusiones erróneas y de esta manera, obtener una ventaja competitiva respecto a rivales”

(Paco González (MoneyBall Data), 2022)

5.3.1 Eventing

Es el conjunto de datos en el que se recogen todas las acciones con el balón, que se llevan a cabo durante un partido. Dependiendo del proveedor el número de eventos que se registran pueden estar entre 1500 y 3500.

Cada evento que se produce tiene que llevar una serie de características como: tipo (si es un pase, un tiro...), resultado (si ese evento ha sido o no realizado con éxito), equipo y jugador por el que ha sido realizado, minuto y segundo y coordenadas (x e y).

El eventing tiene como finalidad, realizar una serie de visualizaciones, que se analizarán antes de visualizar un vídeo del partido o un tramo de partido.

Esas visualizaciones son las siguientes:

1. Mapas de acción de un jugador.

Son mapas de calor en los cuales se puede visualizar en que zonas ha realizado una acción un jugador durante los 90 minutos. Ver Imagen 11.

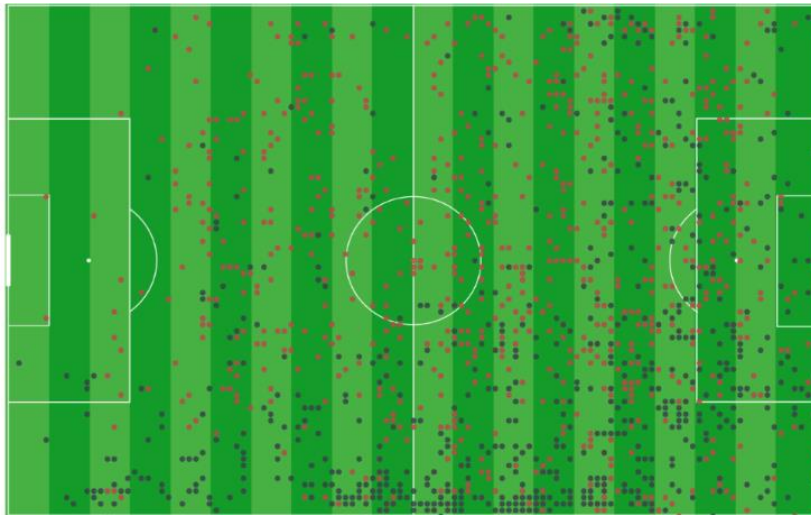


Imagen 11. Mapa de acción de Mbappe y Griezmann en un partido con su selección. Fuente: (Objetivo analista)

2. Distribución de acciones de cada jugador dentro del campo

Se pueden filtrar, dependiendo del tipo de información que queramos obtener. Ver Imagen 12.



Imagen 12. Distribución de acciones Partido Real Madrid – Barcelona. Fuente: (MoneyBall Data)

3. Mapa de pases

Se distribuyen por el terreno de juego los diferentes pases a lo largo del partido. Se diferencia la posición de inicio y la posición final del pase. Además, mediante colores, podemos diferenciar si el pase es acertado o no. Ver Imagen 13.

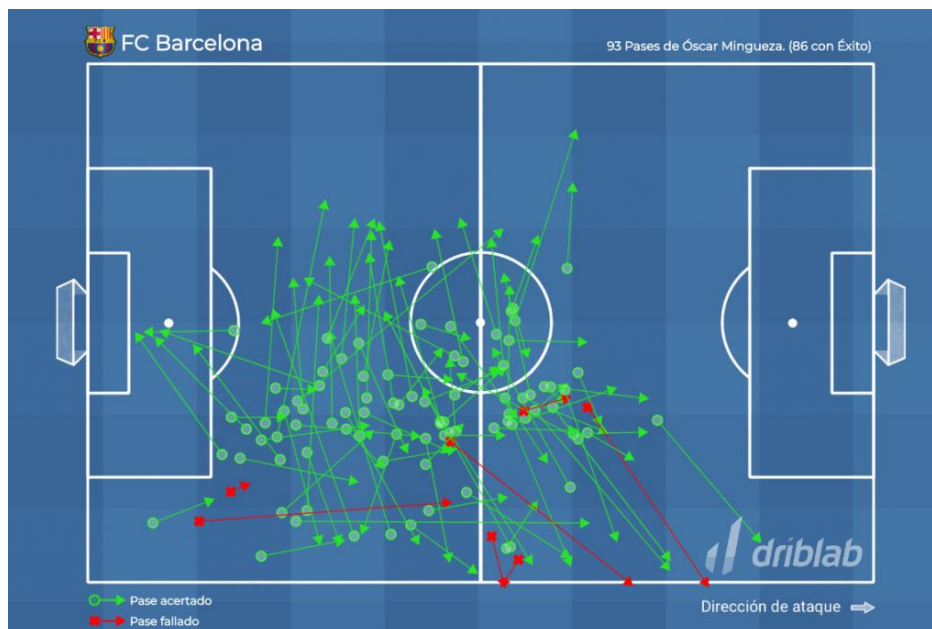


Imagen 13. Distribución de pases Mingueza. Fuente: (Driblab)

4. Simulación de xPTS

A partir de diferentes eventos acontecidos, se calcula los puntos esperados por cada equipo en un partido. Ver Imagen 14.

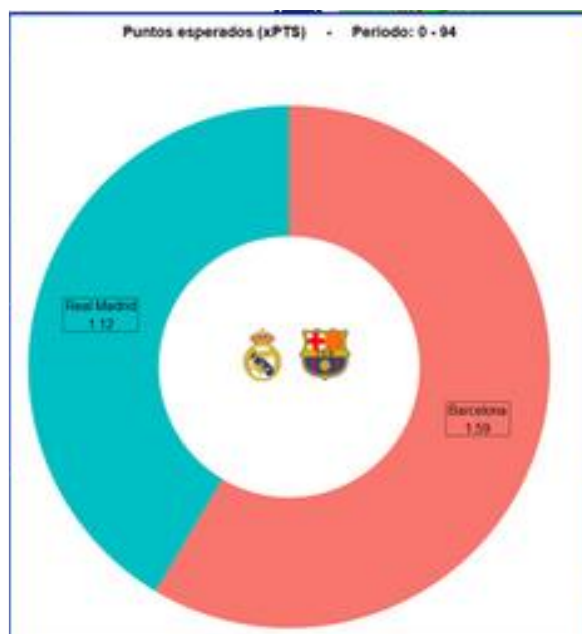


Imagen 14. Puntos esperados en el Partido Real Madrid – Barcelona. Fuente: (MoneyBall Data)

5. Porcentaje de acciones en cada zona del terreno de juego

Visualizamos el % de acciones en cada parte del terreno de juego con respecto al total realizadas en un partido. Ver Imagen 15.

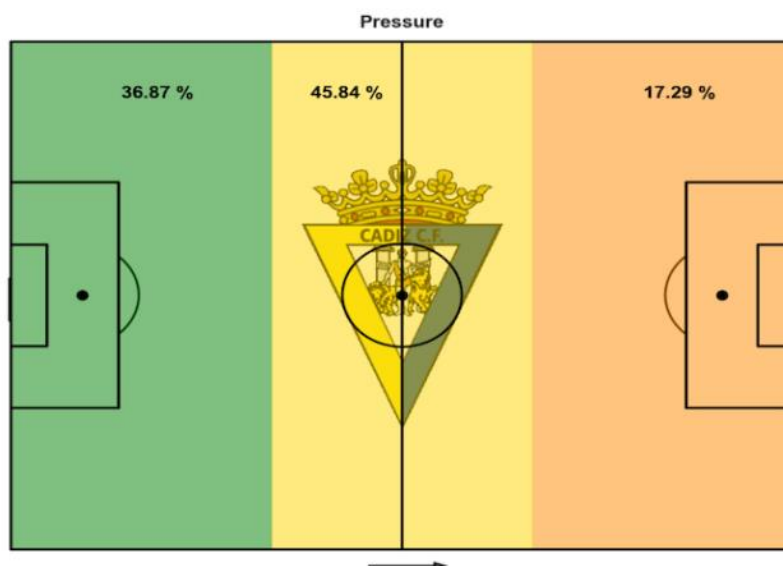


Imagen 15. % de presión realizada en cada zona del campo por el Cádiz CF. Fuente: (Marca)

6. Disparos (visiones frontal y panorámica)

Se representan los disparos en diferentes gráficos, una en visión frontal, y otra en visión panorámica. Ver Imagen 16, Imagen 17.

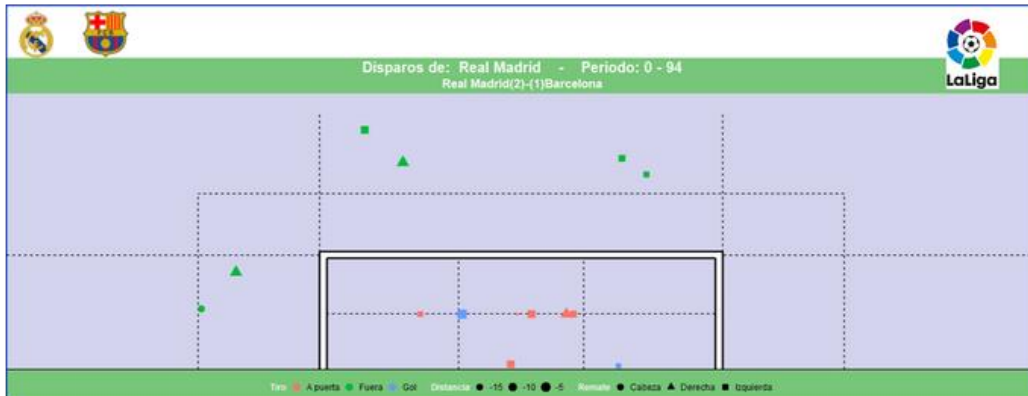


Imagen 16. Mapa de disparos en visión frontal partido Real Madrid – Barcelona. Fuente: (MoneyBall Data)

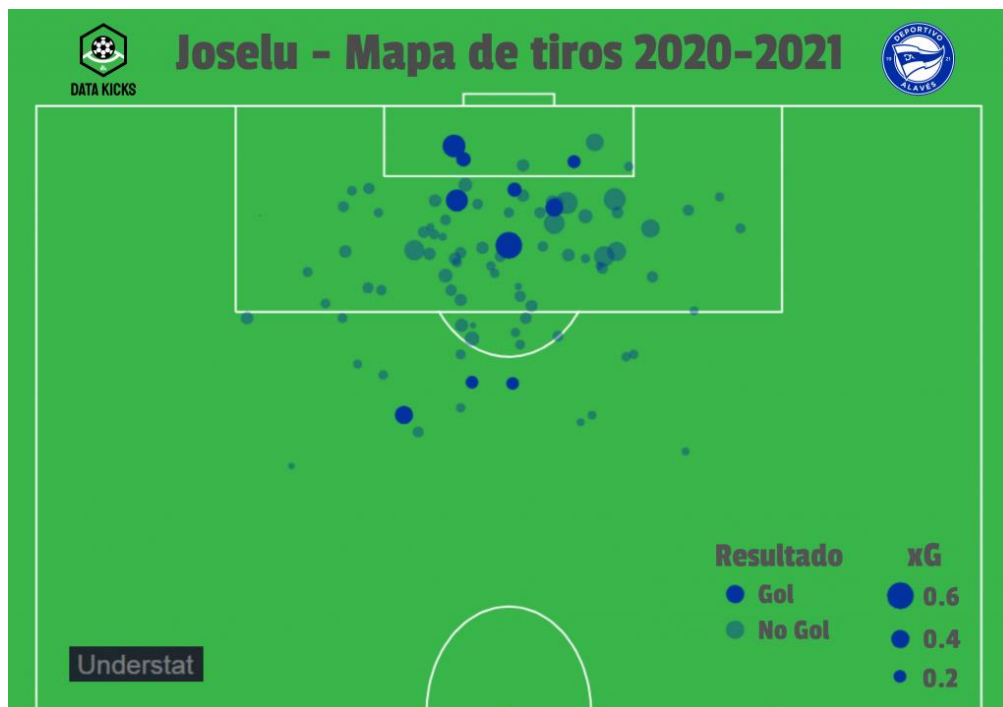


Imagen 17. Mapa de tiros visión panorámica. Joselu temporada 2020/2021. Fuente: Data Kicks

5.3.2 Tracking

Es un conjunto de datos que recoge la posición de cada jugador en cada instante de juego. A diferencia del eventing, este, no recoge acciones determinadas, como

pueden ser tiros, goles, pases... Se recogen datos con una frecuencia de 25 frames (fotogramas) por segundo.

A través del tracking, podemos realizar las siguientes visualizaciones:

1. Mapa de calor

Representación a través de colores, de las zonas donde se ha posicionado un jugador/equipo en los 90 minutos. Ver Imagen 18.

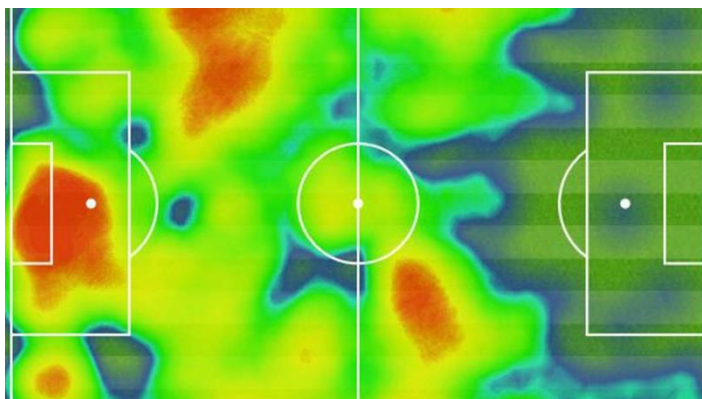


Imagen 18. Mapa de calor Espanyol (Partido Real Madrid – Espanyol) Fuente: (Marca)

2. Diagramas de Voronoi

Nos permite analizar el área de influencia de cada jugador en un instante, diferenciando los componentes de cada equipo por color. Ver Imagen 19.

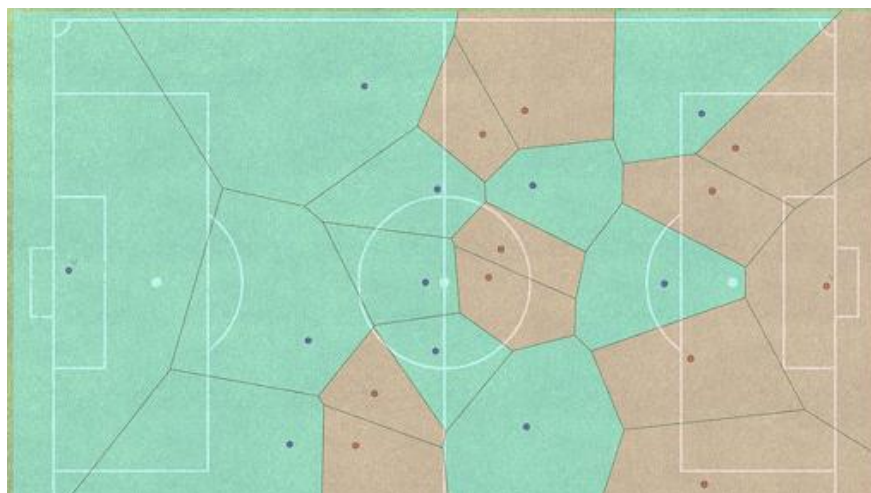


Imagen 19. Diagramas de voronoi. Fuente: (Objetivo Analista)

3. Gráfico de barras de parámetros físicos

Vemos la comparativa de un jugador con respecto al resto de jugadores.
Ver Imagen 20.

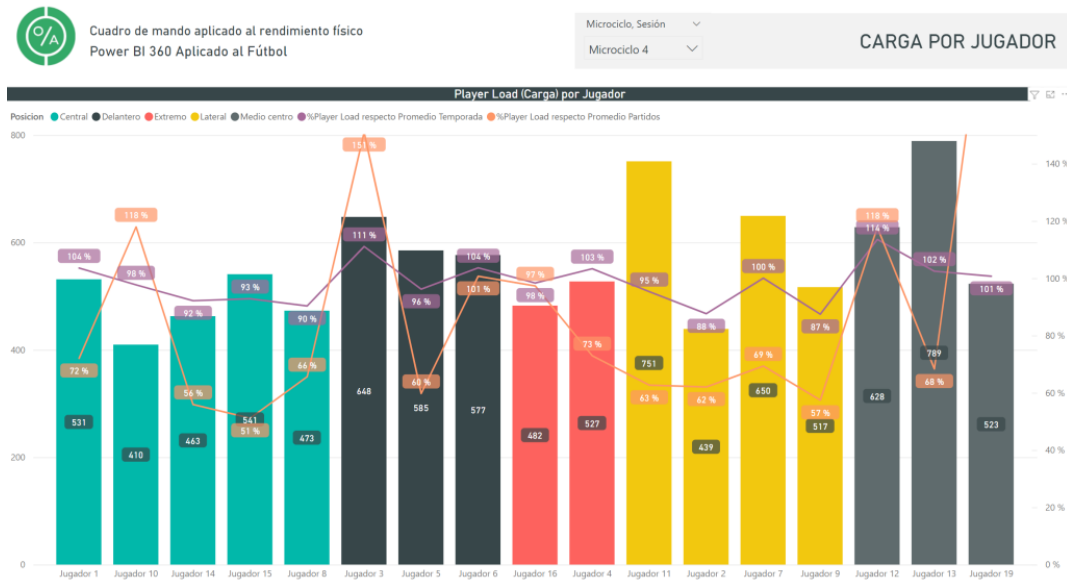


Imagen 20. Carga por jugador. Fuente: (Objetivo Analista)

5.4 Wellness control y prevención de lesiones

5.4.1 Introducción

El control de carga hoy en día es uno de los procesos fundamentales para optimizar el rendimiento de los diferentes jugadores. Se evalúa la fatiga para minimizar el posible riesgo de lesión. Estos factores internos van a servir para evaluar si un futbolista está respondiendo de manera adecuada a la carga de entrenamientos.

Con los diversos sistemas como GPS, pulsómetros o medidores de lactato (sistema para medir la glucosa dentro de la sangre) se podrían medir de manera precisa todos estos aspectos de salud del jugador. Pero, no siempre se tiene la posibilidad de tener todos estos instrumentos, ya que un club de primera división tiene acceso a ellos, pero un club de 2 RFEF no. Lo ideal es reunirse diariamente con cada jugador para hablar de cómo se encuentra, pero debido a la alta carga de partidos esto resulta inviable. Ahí es donde entran los cuestionarios wellness. (Imagen 21)

Estos cuestionarios virtuales representan lo que piensa cada futbolista de su propio estrés y dan información tanto a nivel físico como a nivel psicológico. A través de estos cuestionarios se alimenta la base de datos.

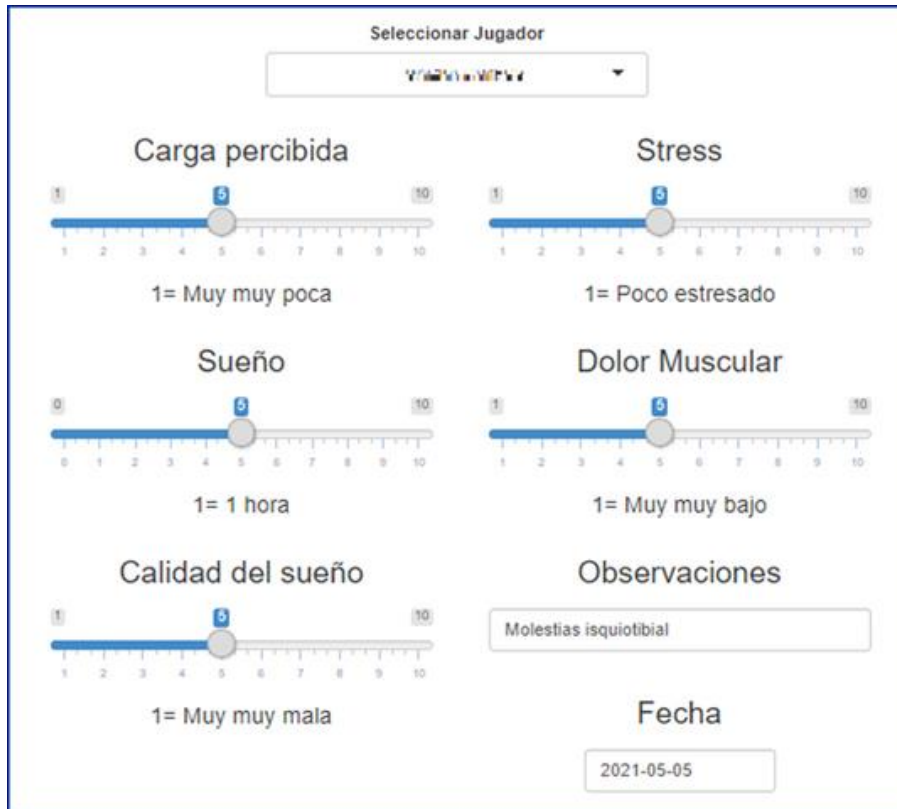


Imagen 21. Cuestionario wellness. Fuente: (MoneyBall Data)

5.4.2 Parámetros clave en la prevención de lesiones

Los preparadores físicos y otros profesionales de las entidades deportivas identifican una serie de parámetros físicos tanto objetivos como subjetivos con el objetivo de minimizar el riesgo de lesiones.

En cada uno de los parámetros se debe visualizar de manera clara la visualización de la demarcación de cada jugador y los valores medios de cada uno de ellos.

1. Distancias y velocidades

Los parámetros que analizar son los siguientes:

-*Distancia total*: Es la suma total de la distancia recorrida por un jugador en un entrenamiento o partido.

-*Distancia promedio*: Es la media de las distancias recorridas por un jugador en cada entrenamiento o partido.

-*Distancia a alta velocidad*: Es la suma de las distancias a alta velocidad (más de 21 km/h)

-*Velocidad máxima*: Velocidad máxima de cada jugador en un entrenamiento o partido.



-*Distancia por minuto*: Se calcula la distancia recorrida por cada minuto para cada jugador.

2. Aceleraciones

Las aceleraciones es uno de los parámetros más importantes a analizar, ya que es donde se produce la mayor fatiga muscular y por tanto la mayor probabilidad de lesión del jugador.

Los parámetros que analizar son los siguientes:

-*Número de aceleraciones*: Se cuenta el número de aceleraciones a más de $3m/s^2$

-*Número de deceleraciones*: Se cuenta el número de deceleraciones a más de $3m/s^2$

3. Carga de trabajo

La carga de trabajo se cuantifica de manera objetiva (a través de los dispositivos GPS) y de manera subjetiva (a través de cámaras y sensaciones del propio jugador).

Por ello, seleccionamos los siguientes parámetros de estudio:

-*Carga de trabajo objetiva frente a la subjetiva*: Se suma la carga de trabajo recogida por los dispositivos GPS frente a la carga subjetiva obtenida con las pruebas wellness explicados anteriormente.

4. Distancia a alta intensidad por minuto

Cuánto más alta sea la intensidad de trabajo, mayor será la probabilidad de lesión. Por tanto, se analiza la distancia a alta intensidad por minuto de cada jugador. Ver Imagen 22.

Distancia con velocidad más de 5.5 m/s (por minuto, en comparación con la media de la distancia de los jugadores del campo), m

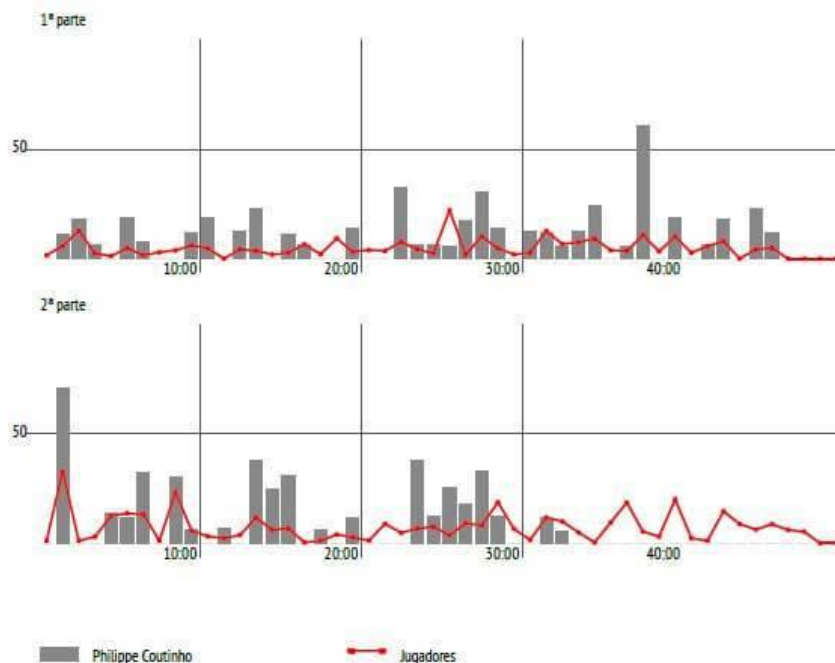


Imagen 22. Estudio de distancia recorrida a alta intensidad Coutinho en comparación con el resto de los jugadores. Fuente: (Objetivo analista)

5.4.3 Wellness control

1. Horas de descanso

Cuando se habla de descanso, quiere decir realizar un reposo absoluto de cualquier actividad. Para garantizar que los procesos de regeneración muscular sean lo más efectivos posible se debe controlar ese sueño. Cuánto mayor es el desgaste físico, mayor serán las horas de descanso. Para futbolistas profesionales estas horas de sueño, deben estar situadas entre 10-12, dependiendo de la constitución de cada año.

2. Estrés

Un jugador que experimenta altos niveles de estrés durante periodos largos y continuos le convierte en convertirse en un jugador más propenso a sufrir una lesión.

3. Fatiga y dolor muscular

Si un músculo entra en un estado de fatiga es más posible que ocurra una lesión muscular. Una vez se identifique esta fatiga, se debe adaptar la carga para evitar una posible lesión.

Estos parámetros se pueden observar en la siguiente Imagen 23.

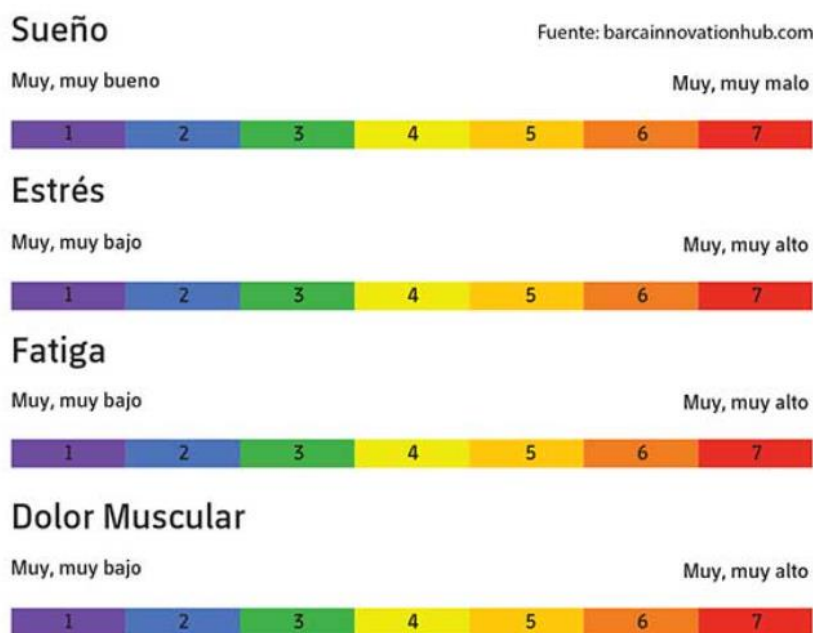


Imagen 23. Test de wellness control. Fuente: (Barcainnovationhub)

5.5 Scouting deportivo y detención del talento

5.5.1 Introducción

“¿Qué es scouting? Es el análisis científico que realiza un Técnico Deportivo recabando la máxima información posible sobre el elemento que se le asigne, utilizando los medios tecnológicos adecuados para transmitirla a los interesados, minimizando, los riesgos que puede causar la competencia”

(Jesús Botello (Analista deportivo), 2018)

El scouting con los tiempos que corren hoy en día, es una de las herramientas más potentes para sacar ventajas con respecto a otros rivales en el mundo del fútbol. Años atrás, era algo más visual y subjetivo, pero ahora se analizan cualquier tipo de parámetros a través de los datos. Antaño, los ojeadores (Imagen 24), scouts o directores deportivos, recorrían los fines de semana estadios de diferentes equipos viendo partidos.



Imagen 24. Ojeador. Fuente: (Diario de Mallorca)

Actualmente, sigue sucediendo esto, pero de una manera diferente.

"Conmigo trabajan 15 personas y dedicamos un cuidado integral a cada competición. La información recopilada se comparte: cada mes, los técnicos hacen un top 11 del torneo que siguen. Esto hasta diciembre, y comenzamos a recopilar muchos jugadores en la base de datos. A partir de ahí aplicamos filtros para tratar de llegar a abril con 120 nombres, 11 o 12 por cada puesto. Luego comenzamos a profundizar sobre los perfiles".

(Ramón Rodríguez (Monchi), 2018)

Ahora la mayoría de las direcciones deportivas trabajan primero con un análisis masivo del dato. Una vez tienen recogidos los datos de todos los jugadores, se encargan de elegir cuales serían los mejores para su equipo en cada posición. Una vez tienen esa lista de nombres, es cuando empiezan a ver partidos en directo de forma masiva, pero ya con objetivos y nombres claros, habiendo sido analizados previamente.

Estas bases de datos de las que hablamos, las podemos dividir en 4 grandes bloques:

1. Datos técnico-tácticos:

Pueden ser a nivel de equipo, nivel individual, dependiendo de lo que se quiera analizar.



2. Datos físicos y de lesiones:

Están relacionados con los técnico - táctico además de ayudarnos al wellness o la prevención de lesiones.

3. Datos subjetivos:

Proviene de los informes por parte de los scouts por el conocimiento del juego.

4. Datos cognitivos:

Son los datos que tienen que ver con cómo está mentalmente el jugador, y su estado de ánimo.

“La valoración en conjunto de estos cuatro pilares fundamentales, nos permitirá conocer si un jugador puede ser adecuado para recomendar su fichaje por parte de la dirección deportiva del club.”

(Paco González (MoneyBall data) 2022)

5.5.2 Datos técnico – tácticos

Son las acciones que realiza un jugador dentro de un partido y que pueden ser analizadas.

Algunos ejemplos de datos técnico – tácticos son los siguientes:

1. Clásicas

Minutos jugados, partidos jugados...

2. Eficiencia

Pases cortos, pases largos, pases a media distancia...

3. De disciplina

Número de tarjetas amarillas/rojas, faltas recibidas, realizadas...

4. Ofensivas y defensivas

Asistencias, goles a favor, goles en contra, tackles, duelos aéreos...

5. Espacio

Zona del campo donde se recogen las acciones dentro de un partido.



5.5.3 Datos físicos

Se recogen todos los datos de los futbolistas que tengan que ver exclusivamente con su físico. Ya sea bien en entrenamientos o bien en partidos. Un ejemplo de ellos pueden ser la velocidad, aceleración máxima, fuerza...

Estos datos, como bien he comentado en apartados anteriores, se recogen mediante los GPS que lleva cada jugador. Todos estos datos físicos, son de gran utilidad, aparte de para el scouting y la detención del talento, para los preparadores físicos, a la hora de poder detectar las posibles lesiones y planificar entrenamientos.

5.5.4 Datos subjetivos

Son los datos que se obtienen de la visualización de partidos y valoraciones que realizan los scouts. Estos informes se hacen de manera específica y cada scout o grupo de trabajo tendrá una metodología u otra dependiendo de la situación. Al igual que con las valoraciones de los jugadores, para un grupo de trabajo pueden ser del 1-5 como vemos en la Tabla 1 y para otros del 1-10 o del 1-100.

Valoración Subjetiva						
Variable / Valor	5	4	3	2	1	Puntuación
Capacidad de recuperación entre esfuerzos de alta intensidad en la 1ª parte	Muy Buena	Buena	Aceptable	Mala	Muy Mala	
Capacidad de recuperación entre esfuerzos de alta intensidad en la 2ª parte	Muy Buena	Buena	Aceptable	Mala	Muy Mala	
Rendimiento en acciones explosivas (sprints, saltos, acciones de 1x1...) en la 1ª parte	Muy bueno. Ganador en la mayoría de acciones de este tipo.	Bueno. Me sentí rápido, explosivo.	Aceptable, pero no en mi mejor nivel de rendimiento.	Malo. Me sentí poco explosivo.	Muy malo. Perdedor en la mayoría de acciones de este tipo.	
Rendimiento en acciones explosivas (sprints, saltos, acciones de 1x1...) en la 2ª parte	Muy bueno. Ganador en la mayoría de acciones de este tipo.	Bueno. Me sentí rápido, explosivo.	Aceptable, pero no en mi mejor nivel de rendimiento.	Malo. Me sentí poco explosivo.	Muy malo. Perdedor en la mayoría de acciones de este tipo.	
Rendimiento general en los últimos 15 minutos del partido	Muy bueno. Me sentí poderoso en los últimos minutos del partido.	Bueno. Me sentí cómodo en los últimos minutos del partido.	Aceptable. Con las reservas justas para terminar el partido.	Mal. Bastante cansado.	Muy mal. Muy cansado. Extenuado.	
					Total	



Tabla 1. Valoración subjetiva. Fuente: (twitter (miguelangelcampos))

Estas valoraciones se pueden obtener de diferentes maneras, mediante la visualización de entrenamientos/partidos en directo o bien por televisión ya sea en directo o en diferido.

Cuántos más informes se tengan de cada jugador, más veracidad tendrá la valoración de la dirección deportiva.



5.5.5 Datos cognitivos

Aprendizaje, toma de decisiones, resolución de problemas...

El estado de ánimo de un jugador lo podemos medir mediante test o aplicaciones mediante las cuales este jugador responde a una serie de preguntas. También un punto a observar son sus redes sociales o preguntando directamente a personas de otros clubs en los que ha jugado como es este jugador.

“David Villa me convenció. Hablé con él, un excompañero del calado de David Villa, cuando hablé con él me marcó mucho. Yo he hecho 600 llamadas de estas, llamando a personas para hablar de excompañeros y hay un 70% que es un juicio tibio. Cuelgas y te has quedado igual. Yo también lo hago cuando me ha llamado gente, te da mucha responsabilidad, eh. Opinar sobre otro club... Si no es un íntimo amigo mío, me cuesta no dar un juicio tibio porque me parece meterte en las declaraciones de otras personas. En el caso de David Villa su juicio no fue tibio. Me dio mucha confianza, fue clave en la contratación de Jack Harrison”.

(Victor Orta (Director deportivo Leeds), 2021)

Existen muchos aspectos que pueden afectar en el estado de ánimo de un jugador y por tanto en su rendimiento. Atención, confianza, ansiedad...

Otro papel muy importante que valorar es el entorno en el cual se encuentra el jugador. Un experto en ello es José Rojo Martín “Pacheta” que en todos sus equipos se reúne en los inicios de temporada con las familias de los futbolistas a explicarles que forman un papel fundamental en la vida del jugador y por tanto en el futuro de la temporada.

“Me gusta transmitir mi idea a todos los que me rodean. Quedé con las mujeres y estuvimos hablando un rato. Ellas son una parte crucial en este proyecto común. Estas cosas de otra índole que hacemos con regularidad, y que no son lo propio del entrenamiento, también tienen su significado. Intentamos cerrar el círculo de todo lo que afecta al futbolista”

(Pacheta (Entrenador Real Valladolid), 2019)

5.5.6 Herramientas de Scouting Deportivo y detención del talento

La creación de una herramienta que una la mayor cantidad de datos posibles es la clave del éxito. Pero hay que tener en cuenta cada situación y cada demanda. Es decir, si nos piden un cuchillo de sierra, no podemos entregar una caja con diferentes tipos de cuchillos, debemos entregar un cuchillo de sierra. Menos, es más.

El uso principal de esta herramienta será la búsqueda de jugadores en función de diferentes parámetros y comparándola con los demás jugadores de la base de datos.

Una vez aplicados todos los filtros, debemos ejecutarlos y que sea la propia herramienta la que nos de los resultados de una manera visual.

Vamos a ver un ejemplo (Imagen 25) de cómo realizaríamos una búsqueda de un jugador que haya marcado más de 10 goles y haya dado más de 5 asistencias.



Imagen 25. Herramienta de scouting deportivo para acciones ofensivas. Fuente (elaboración propia)

Con este ejemplo se puede comprobar que a nuestra dirección deportiva le estamos dando ese cuchillo que nos ha pedido. Nos pide jugadores con más de 10 goles y 5 asistencias. Le mostramos todos los que cumplan esos requisitos además de mostrarle gráficas de manera ilustrativa para una mejor comprensión.





6. CASO PRÁCTICO

6.1 Introducción

Para un mejor desarrollo del caso práctico y así ver como se implementa una herramienta de Scouting Deportivo dentro del mundo real mediante Power Bi, vamos a analizar y estudiar un caso real.

Vamos a suponer que formamos parte del equipo de trabajo de la dirección deportiva del Borussia Dortmund. Ante todos los rumores que han surgido desde verano en torno al interés por uno de sus jugadores más importantes y una de las mayores perlas a nivel mundial, Jude Bellingham, por diversos clubes importantes (Real Madrid, Manchester United, Chelsea...). El director deportivo Sebastian Kehl, nos manda que busquemos sus posibles 3 mejores sustitutos en cuánto a lo puramente estadístico (datos).

Para ello vamos a realizar un estudio completo del caso y finalmente mostraremos una solución lo más clara y precisa de forma gráfica con la herramienta Microsoft Power Bi.

Esta herramienta va a centrarse específicamente en la sustitución de nuestro jugador Jude Bellingham, por ello, tanto todos los parámetros como todas las visualizaciones van a servir para compararlo a él con sus posibles sustitutos, no va a ser una herramienta de búsqueda de jugadores a nivel global, como podría ser la ilustrada en la Imagen 25.

Cabe destacar que en la realidad de los clubes profesionales el dato sólo sirve de ayuda. Hay un procedimiento muy complejo durante todo un año por parte de la secretaría técnica y scouts para el fichaje de un jugador. Se realiza un seguimiento en bruto y en neto de todos los jugadores de ligas similares o de ligas que puede permitirse ese fichaje económicamente. Una vez se ha realizado este seguimiento se da una lista de los mejores jugadores por parte de los scouts y analistas a la dirección deportiva y esta es la que elige. Nunca se va a fichar a un jugador sólo por los datos (forma únicamente objetiva) y nunca se va a fichar a un jugador por sólo mirar como juega y la intuición (forma únicamente subjetiva).

6.2 Identificación del cliente

Lo primero a determinar antes de embarcarnos en cualquier tipo de proyecto, sea del mundo empresarial, deportivo... es conocer bien para quien realizamos este proyecto.

Por ello vamos a estudiar primero a la entidad y luego a las personas para las cuales trabajamos, que son el entrenador y el director deportivo.

6.2.1 Borussia Dortmund

El Borussia Dortmund es un club alemán con sede en Dortmund, más concretamente en Westfalia, Alemania. Es un club centenario, ya que fue fundado en 1909 y es actualmente, el segundo equipo más laureado del país alemán, después del todopoderoso Bayern de Múnich.

Su palmarés también es bastante amplio. Cuenta con 8 Bundesliga, 5 copas de Alemania y 6 supercopas alemanas, esto a nivel nacional. A nivel internacional cuenta con 1 Recopa de Europa (1966), 1 Liga de Campeones y 1 copa Intercontinental.



Imagen 26. Signal Iduna Park. Estadio Borussia Dortmund. Fuente: (Bundesliga)

En la Imagen 26, podemos ver el Signal Iduna Park, actual estadio donde juega sus partidos el Borussia Dortmund. Tiene capacidad para 81365 espectadores y está catalogado como uno de los estadios de máxima categoría de la UEFA. Es el estadio más grande de Alemania, seguido por el Allianz Arena, estadio del Bayern de Múnich.

6.2.2 Sebastian Kehl

Sebastian Kehl es el director deportivo (Imagen 27) del Borussia Dortmund. Nació en Hessen, Alemania, el 13 de Febrero de 1980.

Antes de desempeñar el cargo de director deportivo tuvo una gran trayectoria como futbolista (Imagen 28). Se desempeñaba como defensa central o centrocampista de carácter defensivo.



Imagen 27. Sebastian Kehl como director deportivo. Fuente: (Twitter)



Imagen 28. Sebastian Kehl como futbolista. Fuente: (Transfermarkt)

A los 16 años, se trasladó de Hessen (su ciudad natal) a Hannover para jugar con las categorías inferiores del Hannover 96, por aquellas en la segunda división alemana.



Con 18 años, debutó con el primer equipo y llegó a jugar hasta 32 partidos. Eso, hizo que varios ojeadores se fijaran en este jugador joven y ese verano acabó fichando por el Friburgo, para jugar en la Bundesliga (Primera división de Alemania).

Después de llevar temporada y media en el club y tras sus grandes apariciones que permitieron que el mismo se pudiera clasificar para disputar la UEFA (actual Europa League) decidió fichar por el Borussia Dortmund, el cual acabaría siendo el club de su vida.

En su primer año en Dortmund, consiguió ganar la Bundesliga y llegar hasta la final de la UEFA, la cual acabó perdiendo contra el Feyenoord de Rotterdam.

Su segundo año en Dortmund fue un poco peculiar. El Bayern de Munich había llegado a un acuerdo con el jugador para que jugase con el club Bávaro las siguientes temporadas. Finalmente, en el verano de 2001, Sebastian decidió quedarse en Dortmund y por tanto rechazar la propuesta del Bayern. Esto hizo que el Bayern amenazase hasta con emprender acciones legales, pero al final por el bien común todo se dejó estar. Esto hizo que esa temporada empezase un tanto extraña para Kehl y quedasen a 17 puntos del Bayern en Bundesliga.

Durante los siguientes años Kehl siguió jugando prácticamente todo y siendo una pieza fundamental para sus entrenadores. En 2008, con la entrada de Jurgen Klopp al banquillo del Dortmund, hizo que Kehl pasase a ser el primer capitán del club.

En 2013 consiguió hacer historia logrando clasificarse para la final del torneo más prestigioso del mundo como es la UEFA Champions League. Final que perdieron contra el Bayern por 2-1.

El 23 de Mayo de 2015 jugó su último partido con el Borussia Dortmund y decidió colgar las botas.

Además de jugar para el Borussia de Dortmund durante la mayoría de su carrera, también jugo para la selección nacional de su País, la selección alemana. Disputó el mundial de 2002 y la Eurocopa de 2004 y jugó hasta 31 partidos con la misma.

Su palmarés consta de 3 Bundesliga, 2 supercopas de Alemania y una copa alemana.

Una vez se retiró, Sebastian empezó a formarse para poder trabajar en un futuro ligado al mundo del fútbol. En 2018 empezó a trabajar para el Borussia Dortmund en diversas funciones dentro del club, y el 1 Julio de 2022 se le nombró nuevo director deportivo del club. Desde ese momento se encarga de la organización de toda la parcela deportiva del mismo.

6.2.3 Edin Terzic

Edin Terzic (Imagen 29) es desde finales de la temporada pasada el entrenador del Borussia de Dortmund.



Imagen 29. Edin Terzic. Fuente: (Getty Images)

Ha estado trabajando para diferentes cuerpos técnicos (como interino del club) desde la época de Jurgen Klopp (2008-2015) hasta el 2022.

Para estudiar como juegan sus equipos y así poder establecer el perfil de jugador que estamos buscando (y poder elaborar las métricas más adelante), no tenemos una larga trayectoria. Por ello, vamos a decidir estudiar su única temporada como primer entrenador, es decir la actual (2022-2023).

Viendo la mayoría de los partidos del Borussia de Dortmund de la presente temporada, el equipo ha jugado con un 1-4-3-3, en el cual Bellingham ha jugado de volante. Su estilo suele ser un estilo atrevido, que le gusta vivir en campo contrario con una presión tras pérdida agresiva y con el objetivo de controlar los partidos mediante la posesión de balón. En cuanto a la salida desde atrás, intentan no rifar el balón y salir jugando desde portero con centrales.

Por ello, el perfil que buscamos es un mediocentro con buen trato de balón y que haya recuperado bastantes balones debido a la presión tras pérdida que realiza el equipo. Por ello podemos concluir que tenemos que buscar un jugador con un perfil parecido al que se marcha, es decir, a Jude Bellingham.



6.3 Petición del cliente (Requerimientos)

Este punto es el más importante de todo el proyecto, ya que, si no definimos bien las directrices de lo que desea nuestro cliente, el resto de los pasos del proyecto van a estar mal. Para comprobar que esto es una realidad, vamos a dar un ejemplo muy conocido dentro del mundo del fútbol.

Cuando se marchó Cristiano Ronaldo del Real Madrid (jugador con más goles en la historia del club y líder del vestuario durante 9 temporadas) se tenía que buscar un sustituto. El cliente que en este caso era el Real Madrid, demandaba un jugador que marcara goles, pero que además pudiese "tirar del carro" cuando viniesen problemas. La dirección deportiva del Real Madrid no entendió bien lo que demandaba el cliente y decidió traer como sustituto a Mariano Díaz. A parte de lo costoso a nivel económico del fichaje, analizando sus datos (que además es el tema que nos incumbe) desde ese momento, ¹Mariano ha jugado 1528 minutos, saliendo 11 veces como titular y marcando 7 goles.

Se puede comprobar que la dirección deportiva no entendió bien lo que demandaba el cliente y Mariano se ha convertido en uno de los fichajes más infructuosos de la historia del club madridista.

En nuestro supuesto, la situación es la siguiente:

Jude Bellingham (Imagen 30) (nacido el 29 de Junio de 2003, en Stourbridge, Inglaterra) juega como mediocentro o pivote defensivo. A los 8 años entró a formar parte de la academia del Birmingham. Tras completar su ciclo formativo y aún en edad de juvenil, con 16 años debutó con el primer equipo. Tras una temporada siendo indiscutible con el Birmingham, jugando en Championship (segunda división inglesa), se oficializó el fichaje por 29 millones de euros por parte del Borussia Dortmund.

Jude, ha logrado muchos récords debido a su juventud (ser el jugador más joven en marcar gol con el Borussia de Dortmund o ser el futbolista inglés más joven en salir como titular en un partido de UEFA Champions League)

Todos los datos que vamos a estudiar van a ser de la temporada 2022/2023, ya que, aunque en el momento del estudio aún no se haya completado la temporada entera, estos datos tienen que ser lo más actualizados posibles para estudiar el rendimiento actual de cada jugador.

¹ Datos a fecha 20/03/2023



Imagen 30. Jude Bellingham. Fuente: (BeSoccer)

Lo primero que debemos tener en cuenta son los datos² del jugador que vamos a sustituir, es decir, Jude Bellingham. En el momento actual, Jude ha jugado 24 partidos, saliendo en 23 de ellos como titular y acumulando 2075 minutos. Ha marcado 4 goles y ha repartido 4 asistencias.

Al ser un jugador que se desempeña en el centro del campo, para tener un informe más preciso, debemos recoger datos un poco menos estándar y más específicos de su rol en el terreno de juego. Por ello añadimos, que ha realizado 1241 pases, los cuales han sido acertados 1004, teniendo un acierto en el pase de 79,4 %. Así mismo el porcentaje de acierto en pases cortos es del 86,3%, el de pases medios de 85,6% y el de pases largos de 71,4%. Además de los 1004 pases completados, ha realizado 71 pases que han conducido a tiro del compañero.

También me parecen datos relevantes para un mediocentro sus acciones defensivas. Por ello, Jude ha interceptado 29 balones y ha realizado 22 despejes. En el 1 contra 1 sólo ha sido regateado en 10 ocasiones.

Además de estos datos vistos de forma un poco más general, luego detallaremos unos indicadores que nos ayudarán a tomar la decisión. También deberemos poner en contexto la búsqueda, ya que un jugador como, por ejemplo, Luka Modric, seguramente

²Datos recogidos en fbref.com a fecha 21/03/2023. A lo largo del trabajo estos datos van a variar, ya que van a jugarse más partidos. Por ello, los datos del resultado final serán los datos recogidos en el momento de su realización.



a priori tenga números parecidos o mejores, pero nunca vendría a nuestro club. También debemos tener en cuenta el presupuesto y la edad del jugador (ya que la política de fichajes del Dortmund es fichar a jugadores jóvenes y crear una plusvalía).

Una vez analizada la situación del jugador, sus datos más relevantes (otros no tan relevantes aparecerán a lo largo del caso práctico) vemos que lo que nos demanda nuestro director deportivo (en este caso el cliente) van a ser jugadores que tengan unos datos parecidos o mejores a los comentados anteriormente.

Vamos a elaborar una lista final de 5 jugadores que en cuanto a datos, sean los mejores posibles sustitutos para Jude Bellginham.

Por tanto, ya tenemos los requerimientos de nuestro proyecto y podemos pasar a la siguiente etapa.

6.4 Presupuesto del Proyecto

El presupuesto para la realización de este proyecto se ha acordado en una reunión de la junta de accionistas del club, junto con el director general y el director deportivo a comienzo de temporada. El presupuesto para la parte de scouting será de 200.000 €.

6.4.1 Estimación de costes por actividad

En este apartado estimamos los costes del proyecto en función de otros proyectos similares en otros clubes y temporadas anteriores.

Amortizaciones: Asignamos una parte del presupuesto para amortizar activos fijos, como equipos, según su vida útil y necesidades empresariales.

Software: Destinamos fondos para adquirir y mantener software que mejore la eficiencia y productividad de tu empresa. Consideramos licencias, actualizaciones y software personalizado.

Sueldos: Debemos tener en cuenta que dentro del equipo de analistas hay 2 personas y en el equipo de scouts hay 6 personas.

Viajes: Estimamos estos costos según la frecuencia y duración de los viajes necesarios.

Alojamiento y comidas: Además de los viajes, nuestros scouts van a necesitar alojarse y comer cuando vayan a realizar el visionado de partidos en directo.



NOMBRE	ACTIVIDAD	COSTE
Amortización	Ordenadores	1.000,00 €
	Pantallas	500,00€
Software	Wyscout	800,00€
	Power BI	240,00€
	Instat	600,00€
Sueldos Analistas		40.000,00€
Sueldos Scouts		90.000,00€
Viajes		5.000,00 €
Alojamiento		12.000,00€
Comidas		2.500,00 €
TOTAL		152.640,00€

Tabla 2. Presupuesto

6.4.2 Coste mensual

Teniendo en cuenta que esto es una estimación, tendremos un 47.360€ de colchón sobre el presupuesto para el proyecto.

Hay que tener en cuenta que el pago de cada actividad puede variar en función del número de partidos a visionar en directo, acuerdos de pago con proveedores, etc. Por ello vamos a realizar una estimación mensual.

Teniendo en cuenta que la duración del proyecto es de 9 meses y aproximando que se va a gastar lo mismo en cada vez, nos quedaría un gasto mensual de 16.960€.

6.5 Proceso de ETL

Antes de empezar con el proceso de ETL (extract, transform, load) vamos a explicar brevemente en que consiste y sus pasos.

Es la extracción de los datos de diferentes orígenes, con el fin de transformarlos en indicadores o parámetros (que sean de utilidad en el proceso de negocio) y cargarlos en un único almacén de datos de destino.

6.5.1 Extracción de los datos

Como se ha explicado antes, hay diferentes proveedores de datos técnico – tácticos, que son los que nos interesan ahora, de los cuáles se pueden extraer (Imagen 9).

En el presente trabajo vamos a trabajar con la página <https://fbref.com/es/> , ya que a mi opinión es la que más datos recoge de forma gratuita.

Estadísticas estándar del jugador 2022-2023 Eredivisie

Compartir y exportar ▲ Ocultar a los no clasificados para las estadísticas de calificación. Desp...

Modificar, exportar y compartir la tabla

Obtener como lib Obtén un enlace directamente a la tabla de esta página

Obtener tabla como CSV (para Excel)

Obtener enlace a la tabla

Acerca de compartir herramientas

[Video: Herramientas de uso compartido de SR y cómo hacerlo](#)

[Video: Consejos y trucos de la tabla de estadísticas](#)

[Data Usage Terms](#)

					Edad	Nacimiento	PJ	Titular	M
6	Luuk Admiraal	 NED	DL	Excelsior	21-030	2002	1	0	
7	Nikolas Agrafiotis	 NED	DL,CC	Excelsior	22-354	2000	12	5	
8	Joost van Aken	 NED	DF	Heerenveen	28-336	1994	15	11	
9	Pedro Alemañ	 ESP	CC,DL	Sparta R'dam	21-024	2002	2	0	
10	Hussein Ali	 SWE	CC,DL	Heerenveen	21-044	2002	5	0	
11	Edson Álvarez	 MEX	CC,DF	Ajax	25-172	1997	26	25	2,
12	Pelle van Amersfoort	 NED	CC,DL	Heerenveen	27-013	1996	13	10	
13	Jamal Amofa	 NED	DF	Go Ahead Eag	24-140	1998	25	21	1,

Ilustración 2. Exportación de datos mediante CSV.

Utilizamos la siguiente secuencia:

Obtener datos -> Texto o CSV -> Seleccionamos

Realizamos este procedimiento con los datos de las ligas que nos interesan y los introducimos como origen de datos en Power BI. Una vez los tengamos todos cargados en el modelo de datos de Power BI, como podemos comprobar en la siguiente Ilustración 3, pasaríamos a la transformación de estos datos mediante el programa Power Query, que es una extensión del propio programa Power BI.



	Column1	Column2	Column3	Column4	Column5
1					
2	RL	Jugador	País	Posc	Equipo
3	1	Rodrigo Abascal	uy URU	DF	Boavista
4	2	Issah Abbas	gh GHA	DLCC	Chaves
5	3	Giorgi Aburjania	ge GEO	CC	Gil Vicente FC
6	4	Elias Achouri	tn TUN	CC	Estoril
7	5	Antonio Adán	es ESP	PO	Sporting CP
8	6	Adriano	br BRA	CCDF	Santa Clara
9	7	João Afonso Crispim	br BRA	CCDF	Marítimo
10	8	Lucas Áfrico	br BRA	DF	Estoril
11	9	Martin Agirregabiria	es ESP	DFDL	Famalicão
12	10	Salvador Agra	pt POR	DLCC	Boavista
13	11	Mohamed Aidara	ci CIV	DF	Vizela
14	12	Sotiris Alexandropoulos	gr GRE	CC	Sporting CP
15	13	Ali Alipour	ir IRN	DLCC	Gil Vicente FC
16	14	Allano	br BRA	DLCC	Santa Clara
17	15	Domingos André Ribeiro Almeida	pt POR	CC	Vitória
18	16	Jordi Almeida	br BRA	PO	Paços
19	17	Alejandro Alvarado Jr.	us USA	DLCC	Vizela
20	18	Pedro Álvaro	pt POR	DF	Estoril
21					

Ilustración 3. Power query.

6.5.2 Transformación y carga de los datos

Este es uno de los puntos más importantes de todo el trabajo. Una vez tenemos un gran volumen de datos, tenemos que saber qué hacer con ellos.

“ Lo realmente importante es saber transformar el dato en información”

Ramón Rodríguez “Monchi” (Sport Data Forum 2023)

Por ello, lo primero que vamos a realizar con nuestra herramienta Power Query, va a ser poner un poco más legible la gran cantidad de datos en nuestra base.

Podemos comprobar la siguiente Ilustración 4, cómo de desordenados están los datos.

	APC Column1	APC Column2	APC Column3	APC Column4	APC Column5
	<ul style="list-style-type: none"> Válido 99 % Error 0 % Vacío < 1 % 	<ul style="list-style-type: none"> Válido 99 % Error 0 % Vacío < 1 % 	<ul style="list-style-type: none"> Válido 99 % Error 0 % Vacío < 1 % 	<ul style="list-style-type: none"> Válido 99 % Error 0 % Vacío < 1 % 	<ul style="list-style-type: none"> Válido 99 % Error 0 % Vacío < 1 %
1					
2	RL	Jugador	País	Posc	Equipo
3	1	Brenden Aaronson	us USA	CCDL	Leeds United
4	2	Paxten Aaronson	us USA	CC	Eint Frankfurt
5	3	James Abankwah	ie IRL	DF	Udinese
6	4	Yunis Abdelhamid	ma MAR	DF	Reims
7	5	Himad Abdelli	fr FRA	CCDL	Angers
8	6	Salis Abdul Samed	gh GHA	CC	Lens
9	7	Laurent Abergel	fr FRA	CC	Lorient
10	8	Oliver Abildgaard	dk DEN	CCDF	Hellas Verona
11	9	Matthis Abline	fr FRA	DL	Auxerre
12	10	Matthis Abline	fr FRA	DLCC	Rennes
13	11	Abner	br BRA	DF	Betis
14	12	Zakaria Aboukhlal	ma MAR	DLCC	Toulouse
15	13	Tammy Abraham	eng ENG	DL	Roma
16	14	Christian Acella	it ITA	CC	Cremonese
17	15	Francesco Acerbi	it ITA	DF	Inter
18	16	Álvaro Aceves	es ESP	PO	Valladolid

Ilustración 4. Datos poco legibles

Para ordenarlos vamos a realizar una serie de cambios o modificaciones dentro de Power Query como pueden ser editar encabezados, nombres o dividir columnas. (Ilustración 5)

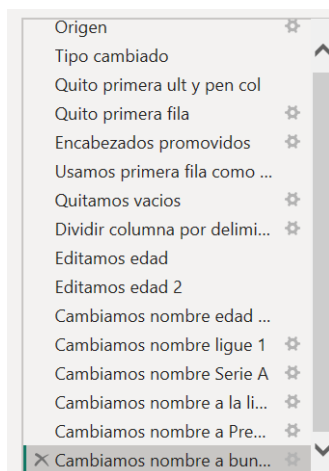


Ilustración 5. Distintos pasos aplicados

El resultado final es el siguiente Ilustración 6.



	Año Jugador	Año País	Año Posc	Año Equipo	Año Comp
	● Válido 100 % ● Error 0 % ● Vacío 0 %	● Válido 100 % ● Error 0 % ● Vacío 0 %	● Válido 100 % ● Error 0 % ● Vacío 0 %	● Válido 100 % ● Error 0 % ● Vacío 0 %	● Válido 100 % ● Error 0 % ● Vacío 0 %
1	Brenden Aaronson	us USA	CCDL	Leeds United	Premier League
2	Paxten Aaronson	us USA	CC	Eint Frankfurt	Bundesliga
3	James Abankwah	ie IRL	DF	Udinese	Serie A
4	Yunis Abdelhamid	ma MAR	DF	Reims	Ligue 1
5	Himad Abdelli	fr FRA	CCDL	Angers	Ligue 1
6	Salis Abdul Samed	gh GHA	CC	Lens	Ligue 1
7	Laurent Abergel	fr FRA	CC	Lorient	Ligue 1
8	Oliver Abildgaard	dk DEN	CCDF	Hellas Verona	Serie A
9	Matthis Abline	fr FRA	DL	Auxerre	Ligue 1
10	Matthis Abline	fr FRA	DLCC	Rennes	Ligue 1
11	Abner	br BRA	DF	Betis	La Liga
12	Zakaria Aboukhlal	ma MAR	DLCC	Toulouse	Ligue 1
13	Tammy Abraham	eng ENG	DL	Roma	Serie A
14	Christian Acella	it ITA	CC	Cremonese	Serie A
15	Francesco Acerbi	it ITA	DF	Inter	Serie A
16	Álvaro Aceves	es ESP	PO	Valladolid	La Liga
17	Mohamed Achi	fr FRA	DL	Nantes	Ligue 1
18	Marcos Acuña	ar ARG	DF	Sevilla	La Liga

Ilustración 6. Datos legibles y ordenados.

Hacemos lo mismo para las diferentes consultas.

Una vez tenemos la gran cantidad de datos legibles y ordenados vamos a ver que datos nos interesan y cuales no.

Los datos que se han cargado son los siguientes:

- Nombre del jugador
- País de nacimiento (nacionalidad)
- Posición
- Equipo
- Competición
- Edad
- Año de nacimiento
- Partidos Jugados
- Partidos jugados como titular
- Minutos disputados
- Numero de minutos jugados totales / 90 (número de partidos jugados del total de los que ha participado)
- Goles
- Asistencias
- Goles + asistencias
- Goles sin penalti
- Goles de penalti
- Numero de penaltis lanzados



- Tarjetas amarillas
- Tarjetas rojas
- Goles esperados incluyendo penaltis
- Goles esperados sin incluir penaltis
- Goles esperados con asistencia
- Goles esperados con asistencia + goles esperados sin incluir penaltis
- Conducciones de balón hacia línea de gol de más de 9 metros
- Pases progresivos (de más de 9 metros) realizados
- Pases progresivos (de más de 9 metros) recibidos
- Resto de datos vistos hasta el momento, pero divididos entre 90 minutos.

Una vez tenemos recogidos todos los datos que nos ofrecen, vamos a eliminar los que no nos van a ser de utilidad para el resultado que deseamos obtener. Por ejemplo, no nos interesan los goles de penaltis marcados por un mediocentro, ya que no nos van a dar información relevante a la hora de fichar. Por ello, vamos a eliminar esta columna.

Utilizamos la siguiente secuencia:

Seleccionamos la columna -> botón derecho -> Quitar

Además, vamos a cambiar el nombre abreviado de las columnas, ya que así nos enteraremos mejor de que dato estamos utilizando.

Doble clic -> Cambiamos nombre

El resultado lo podemos comprobar en la siguiente Ilustración 7.

	AB_C Equipo	AB_C Edad	AB_C Nacimiento	AB_C Partidos jugados	AB_C Titular
	<ul style="list-style-type: none"> ● Válido 100 % ● Error 0 % ● Vacío 0 % 	<ul style="list-style-type: none"> ● Válido 100 % ● Error 0 % ● Vacío 0 % 	<ul style="list-style-type: none"> ● Válido 100 % ● Error 0 % ● Vacío 0 % 	<ul style="list-style-type: none"> ● Válido 100 % ● Error 0 % ● Vacío 0 % 	<ul style="list-style-type: none"> ● Válido 100 % ● Error 0 % ● Vacío 0 %
1	Atl Paranaense	21	2000	28	25
2	Botafogo (RJ)	23	1998	17	16
3	Corinthians	21	2000	25	13
4	Atl Goianiense	22	1999	33	22
5	América (MG)	19	2002	6	0
6	Corinthians	20	2001	20	15
7	América (MG)	31	1990	28	24
8	Internacional	23	1998	34	22
9	Bragantino	22	2000	20	11
10	Cuiabá	22	1999	21	11
11	Santos	22	1999	4	4
12	Fortaleza	22	1999	12	11
13	Coritiba	22	1999	21	16
14	Fluminense	18	2003	3	2
15	São Paulo	28	1993	15	10
16	Atlético Mineiro	24	1997	33	31
17	Bragantino	29	1993	2	1
18	América (MG)	30	1991	28	21

Ilustración 7. Cambio de título y eliminación de parámetros

Además, para luego poder anexar varias consultas, debemos añadir una columna con el nombre de la liga para todas las que no son 5 grandes ligas, ya que, en ellas, ya viene incluida cuando recogemos estos datos.

Columna personalizada -> Añadimos fórmula de columna personalizada -> Aceptar

Esto lo podemos ver en la siguiente Ilustración 8.

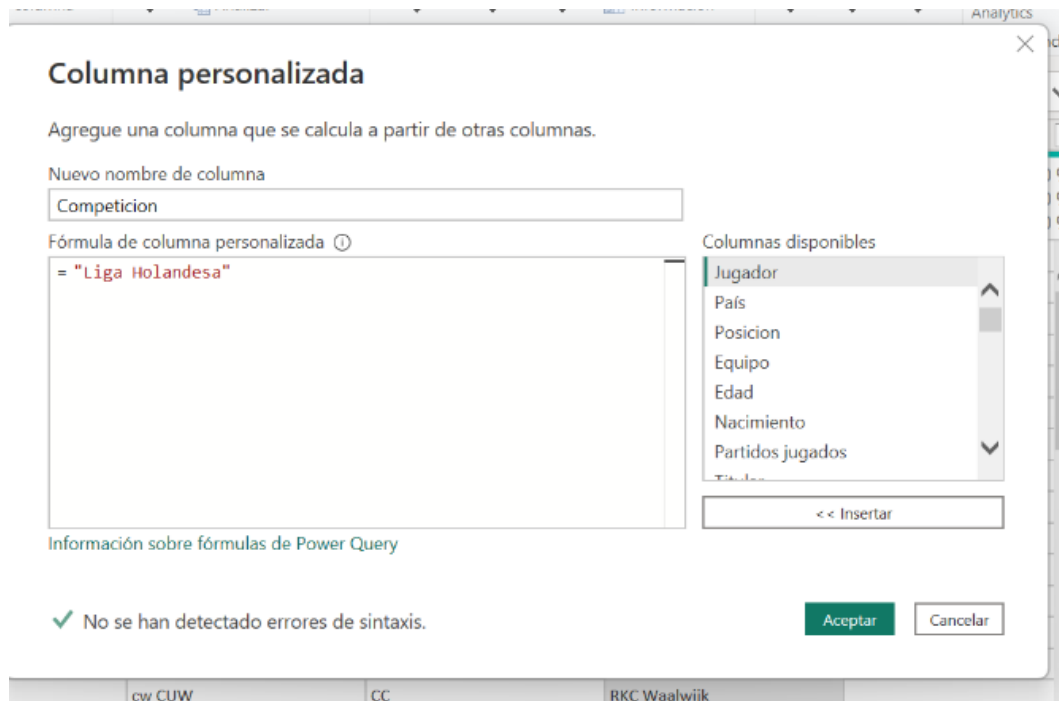


Ilustración 8. Columna personalizada

Por el formato cogido de la página web FBref, lo que ellos ponen como “.” Nosotros queremos ponerlo como “,” luego esto también lo debemos cambiar.

Seleccionamos la columna -> Pulsamos botón derecho -> Reemplazar valores -> Cambiamos el “.” Por la “,”.

Ya solo faltaría cambiar el formato de cada una de las columnas, ya que para los posteriores cálculos debemos de tener los números decimales en formato decimal, los números enteros en formato entero...

Seleccionamos la columna -> Pulsamos botón derecho -> Cambiar tipo -> Seleccionamos el tipo de formato.

Una vez tenemos recogidos y ordenados todos nuestros datos en la base vamos a pasar a detallar que son cada uno de ellos.



6.6 Definición de KPI's o indicadores

Lo primero que debemos hacer es dar una definición o explicar a que nos referimos con KPI's o indicadores. KPI viene del inglés y significa Key Performance Indicator. Los utilizamos como métricas que nos puedan ayudar a la toma de decisiones de un proyecto o proceso. Se pueden utilizar en cualquier área de negocio y por supuesto también en el ámbito del fútbol.

Unos buenos indicadores van a ser diferenciales con respecto al resto de competidores del mismo ámbito, veamos con un ejemplo. A la hora de fichar a un delantero centro (para no hablar siempre del mediocentro) no va a ser lo mismo el club que se fije en el número de goles que marca (primer KPI) que otro club que se fije en el xG (expected goal) (segundo KPI). El número de goles marcados, únicamente nos va a dar un dato de forma cuantitativa, mientras que el xG utiliza una serie de algoritmos predictivos para predecir la capacidad que va a tener un delantero centro de anotar un gol.

Si un jugador tiene un xG más alta que los goles reales anotados, es probable que sea el resultado de malas definiciones o de mala suerte. De la misma manera, si un jugador está marcando más goles que lo que indican sus estadísticas xG, podría ser debido a su calidad.

En pocas palabras, se pueden considerar a las estadísticas xG como una evaluación de las "posibilidades". Mientras que estadísticas de goles anotados únicamente nos va a dar información de los goles anotados en un periodo de tiempo y no de lo que genera ese jugador de manera global.

A continuación, vamos a detallar algunos indicadores que vamos a utilizar en nuestro proceso.

Indicadores generales o básicos:

Son aquellos en los que nos vamos a basar para descartar a la gran mayoría de jugadores que tenemos en nuestra base de datos.

1. Posición del jugador.

Es el primer indicador con el que tenemos que filtrar. De nada sirve querer fichar a un mediocentro y analizar datos de defensas centrales.

Posición del jugador	
Definición	Nos indica en que posición del campo juega el jugador.
Cálculo	No se considera
Excepciones	No se considera



Consideraciones adicionales	Con este indicador ya eliminamos a todos los jugadores que no jueguen en la demarcación del mediocentro.
-----------------------------	--

Tabla 3. Posición del jugador. Fuente: (Elaboración propia)

2. Edad

Con este indicador podemos seleccionar la política de fichaje que queremos. Para nuestro caso, queremos un jugador joven para poder generar más adelante una plusvalía.

Edad	
Definición	Nos indica que edad tiene el jugador.
Cálculo	Año actual – año nacimiento jugador
Excepciones	No se considera
Consideraciones adicionales	No se considera

Tabla 4. Edad Fuente: (Elaboración propia)

3. Minutos jugados

Es otro indicador muy importante para tener en cuenta. Para fichar a un jugador precisamos que ese jugador haya jugado hasta la fecha una serie de minutos, ya que queremos que a su llegada nos dé un rendimiento inmediato, y si lleva parado un tiempo o no ha sido importante para su club anterior, descartaremos su fichaje.

Minutos jugados	
Definición	Nos indica si ese jugador es importante para su club de precedencia o no.
Cálculo	No se considera
Excepciones	No se considera
Consideraciones adicionales	Este indicador es importante ya que no va a ser lo mismo un jugador que haya jugado 1000 minutos y haya marcado 2 goles, que uno que haya jugado el doble y haya metido los mismos.

Tabla 5. Minutos jugados. Fuente: (Elaboración propia)



Indicadores específicos para la posición de mediocentro:

Para el presente estudio vamos a trabajar con los indicadores cada 90 minutos, ya que aparte de que no ha finalizado aún la temporada, creo que es bastante más sencillo de comparar e interpretar los resultados.

- Acciones que intervienen de manera directa en un gol

1. Número de asistencias/partido

No es lo mismo contabilizar las asistencias que da Sergio Busquets o que da Kevin De Bruyne, porque lógicamente, aunque los dos son mediocentros, uno vive más cerca de área rival que otro. Por ello creo que para el perfil de jugador que buscamos como es Jude Bellingham si me parece un indicador de rendimiento propio para el mismo, aunque he de decir que no de los más importantes.

Número de asistencias/partido	
Definición	Nos indica el número de asistencias que realiza.
Cálculo	N de asistencias/90 minutos
Excepciones	No se considera
Consideraciones adicionales	Con este indicador eliminamos jugadores que juegan como pivotes o que tienen un rol más puramente defensivo.

Tabla 6. Número de asistencias/partido Fuente: (Elaboración propia)

2. Non penalty expected goals/partido (npxG/90)

El npxG es muy fácil de explicar. Son los goles esperados sin tener en cuenta los penaltis. Este xG es la probabilidad de que un tiro acabe en gol en función de las características de ese tiro y de los eventos que lo precedieron.

Algunos factores que intervienen en su cálculo son por ejemplo la ubicación del tiro (distancia y ángulo), tipo de pase (centro, balón parado...) tiros con características similares de ese jugador etc....

Este tipo de dato, como ya he explicado anteriormente nos da una mayor información de la que por ejemplo nos daría únicamente el número de goles marcados.

npxG /partido	
Definición	Nos indica la probabilidad de que ese jugador marque un gol cada partido.



Cálculo	$npxG/90$
Excepciones	No se considera
Consideraciones adicionales	Con este indicador eliminamos jugadores que no se atreven a disparar a portería.

Tabla 7. $npxG$ / partido Fuente: (Elaboración propia)

3. Expected assisted goals/partido ($xAG/90$)

xAG o goles esperados con asistencia. Este indicador lo utilizamos para saber la capacidad que tiene un jugador para crear ocasiones de gol sin necesidad de depender del resultado real del tiro. Solo se contabilizan los xAG cuando se realiza un tiro tras un pase completado.

xAG /partido	
Definición	Nos indica la capacidad que tiene un jugador para crear ocasiones de gol
Cálculo	$xAG/90$
Excepciones	No se considera
Consideraciones adicionales	Con este indicador eliminamos jugadores que juegan como pivotes o que tienen un rol más puramente defensivo.

Tabla 8. xAG / partido Fuente: (Elaboración propia)

También vamos a utilizar otro indicador que va a ser $npxG+xAG/90$ con el que se van a indicar las participaciones en goles totales, que únicamente se calculará sumando los valores de $npxG$ y xAG .

Para los siguientes indicadores, vamos a tener que crear una columna y proceder a realizar los cálculos correspondientes, ya que son datos que no nos suministra únicamente FBref.

Seguimos la siguiente secuencia:

Agregar columna -> Columna personalizada -> Escribimos la fórmula a partir de la cual vamos a calcular esa nueva columna de datos -> Aceptar

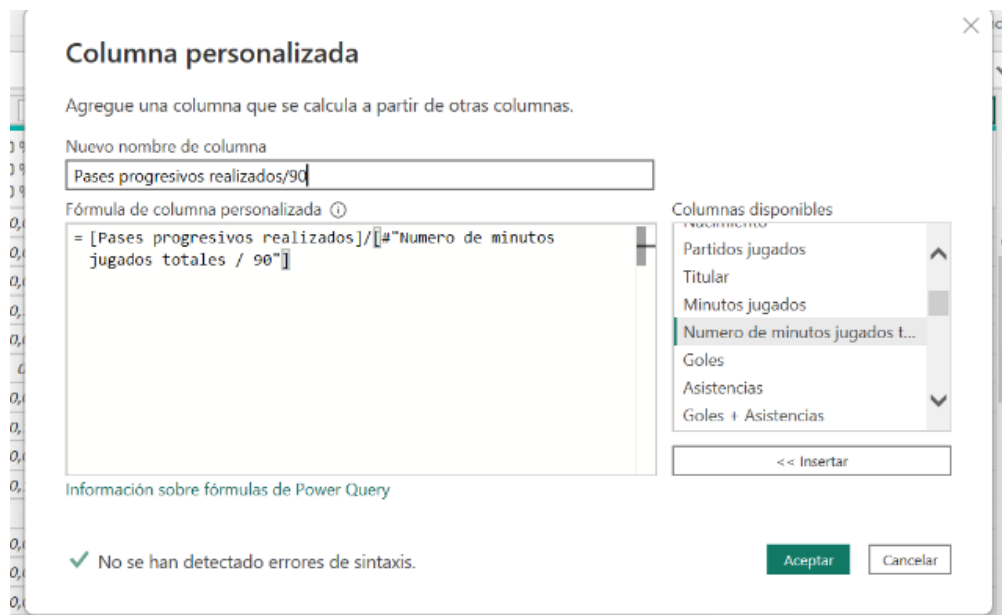


Ilustración 9. Creación de nuevas métricas.

En la Ilustración 9 podemos observar que para tener una mayor precisión, tenemos que dividir a los pases o conducciones en vez de entre el número de partidos jugados, entre el número de minutos jugados en función de los minutos ya que si no, contabilizaríamos igual a un jugador que en un partido ha jugado 5 minutos que a otro que ha jugado 90. De ahí la importancia de minimizar el error lo máximo posible.

Al realizar estos cálculos, las columnas nos van a quedar con muchos decimales y la mayoría de ellos periódicos. Por ello vamos a redondear a 2 decimales como podemos observar en la siguiente Ilustración 10.

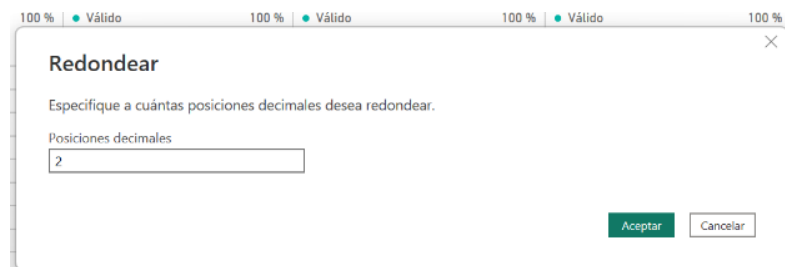


Ilustración 10. Redondeo decimal.



- **Acciones con balón en zona de medios**

1. Conducciones progresivas/partido

Esta métrica la vamos a utilizar para saber el número de líneas de presión que rompe un futbolista en conducción de balón.

Conducciones progresivas/partido	
Definición	Nos indica la capacidad que tiene un jugador para romper líneas rivales
Cálculo	Número de conducciones progresivas totales/ número de partidos disputados
Excepciones	No se considera
Consideraciones adicionales	Con este indicador eliminamos jugadores que sean más posicionales

Tabla 9. Conducciones progresivas/partido. Fuente: (Elaboración propia)

2. Pases progresivos realizados/partido

Esta métrica la vamos a utilizar para saber el número de pases filtrados que realiza un jugador cada partido para romper una línea rival.

Pases progresivos realizados/partido	
Definición	Nos indica la capacidad que tiene un jugador para filtrar balones entre líneas rivales
Cálculo	Número de pases progresivos realizados/ número de partidos disputados
Excepciones	No se considera
Consideraciones adicionales	Con este indicador eliminamos jugadores que den pases únicamente de manera horizontal

Tabla 10. Pases progresivos/partido. Fuente: (Elaboración propia)

3. Pases progresivos recibidos/ partido

Esta métrica la vamos a utilizar para saber el número de pases que ha recibido un jugador entre líneas rivales.

Pases progresivos recibidos/partido	
Definición	Nos indica la capacidad que tiene un jugador para recibir balones entre líneas



Cálculo	Número de pases progresivos recibidos/ número de partidos disputados
Excepciones	No se considera
Consideraciones adicionales	Con este indicador eliminamos jugadores que no se sitúen entre líneas y únicamente reciban pases en posiciones sin transcendencia en el juego

Tabla 11. Pases progresivos recibidos/partido. Fuente: (Elaboración propia)

- Acciones defensivas

1. Recuperaciones de balón/partido

Esta métrica la vamos a utilizar para saber el número de recuperaciones de balón en cada partido.

Recuperaciones de balón /90 minutos	
Definición	Nos indica la capacidad que tiene un jugador para recibir balones entre líneas
Cálculo	Número de recuperaciones de balón/ número de partidos disputados
Excepciones	No se considera
Consideraciones adicionales	Con este indicador eliminamos jugadores que no tengan capacidad de recuperar la posesión de balón

Tabla 12. Recuperaciones de balón/ partido. Fuente: (Elaboración propia)

2. Porcentaje de duelos aéreos ganados

Esta métrica la vamos a utilizar para saber el número de duelos aéreos que gana un jugador a lo largo de la temporada. No vamos a analizar esta métrica por partido ya que no nos daría ninguna información, porque hay partidos que por el modelo de juego del rival hay menos balones aéreos que en otros, por ello queremos tener una métrica más global.



% de duelos aéreos ganados	
Definición	Nos indica la capacidad que tiene un jugador para ganar duelos aéreos
Cálculo	(Número de duelos aéreos ganados/ número de duelos aéreos disputados) x100
Excepciones	No se considera
Consideraciones adicionales	Con este indicador eliminamos jugadores que no ganen duelos aéreos.

Tabla 13. % de duelos aéreos ganados. Fuente: (Elaboración propia)

Una vez tenemos todos los KPI's que queremos utilizar, vamos a juntar las diferentes consultas que tenemos en Power Query en una sola, para que sea mucho más fácil operar con ella más adelante. Para ello vamos a utilizar el siguiente comando:

Combinar/Anexar consultas -> Combinar consultas para crear una nueva -> Aceptar

Esto lo podemos observar en la siguiente Ilustración 11.

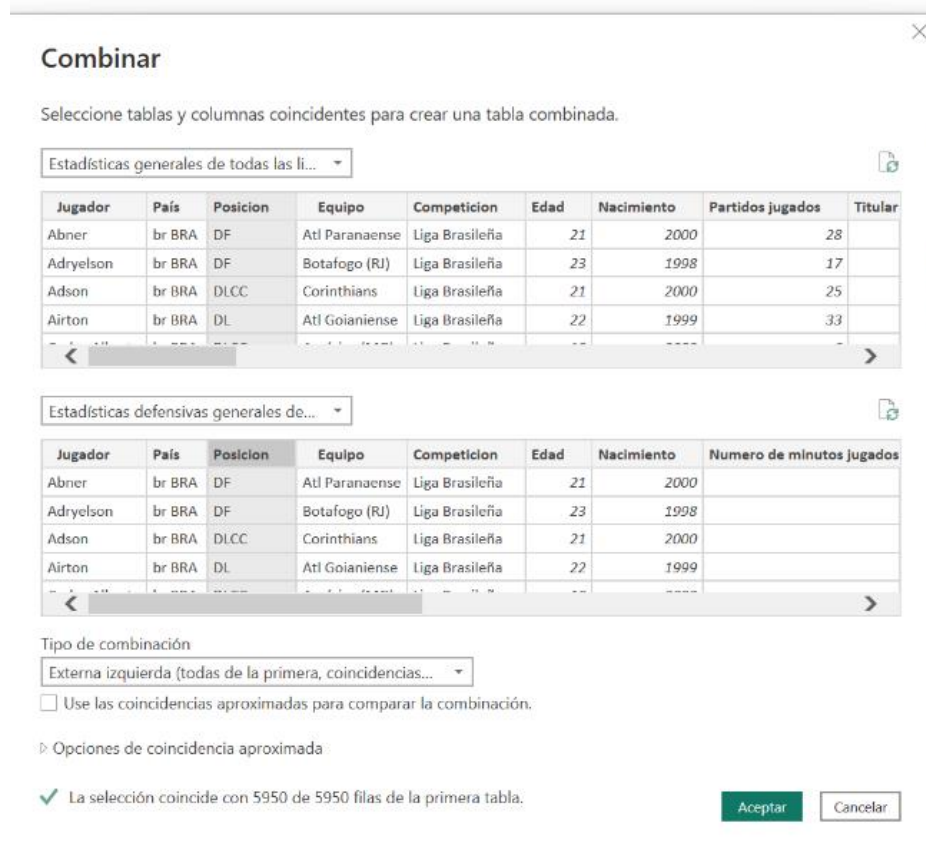


Ilustración 11. Combinar consultas.

6.7 Creación de algoritmo de similitud

Como hemos planteado anteriormente, necesitamos un jugador recambio para Jude Bellingham. Vamos a crear un algoritmo de similitud para saber que jugadores se asemejan a él siguiendo los diferentes indicadores que hemos estudiado.

La aplicación de un algoritmo de similitud permite ampliar nuestra base de datos con información de jugadores con perfiles similares dando a cada indicador de rendimiento un peso, según la importancia que queramos dar al mismo.

6.7.1 Asignación de pesos a las variables

Lo primero que debemos hacer es plantear el perfil de jugador que queremos y a que variables les queremos dar más importancia. Esto es un punto complicado ya que debemos tener en cuenta además del perfil, el estilo de juego del equipo. No vamos a ponderar las mismas variables si queremos saber un sustituto de Sergio Busquets para el FC Barcelona (tendremos que dar más importancia a su juego con balón que sin él, ya que el estilo de juego del Barcelona es un estilo de juego basado en la posesión de este) que un sustituto de Casemiro para el Real Madrid (tendremos que dar más importancia a las recuperaciones de balón ya que con 2 volantes creativos a los costados es el encargado de mantener compacto al mediocampo).



Según el tipo de perfil que hemos detallado en apartados anteriores vamos a asignar una serie de pesos (1 menor importancia - 10 mayor importancia) a las variables:

1. Número de asistencias/partido

No es un dato de relativa importancia, ya que el número de asistencias de gol depende de factores como el compañero que recibe el balón y la habilidad que tenga este de marcar gol o no.

Vamos a ponderarlo con un 3.

2. Non penalty expected goals/partido (npxG/90)

Queremos que nuestro jugador intervenga en los goles del equipo, y sea un mediocentro llegador, pero tampoco es un factor muy importante teniendo en cuenta la posición de este (no buscamos un goleador)

Vamos a ponderarlo con un 5.

3. Expected assisted goals/partido (xAG/90)

Queremos que nuestro jugador tenga mucha relación con las asistencias que da a sus compañeros, independientemente del resultado de gol o no. Por ello creo que este indicador debe ser el que más importancia debe tener.

Vamos a ponderarlo con un 10.

4. Conducciones progresivas/partido

Queremos que nuestro jugador rompa líneas de presión del rival en conducción, ya que es una buena forma de generar hombres libres.

Vamos a ponderarlo con un 7.

5. Pases progresivos realizados/partido

Queremos que nuestro jugador filtre balones para conseguir lo mismo que en el punto anterior. Vamos a darle un poco más importancia que al anterior, ya que me gusta que se mueva el balón y no el jugador.

Vamos a ponderarlo con un 8.

6. Pases progresivos recibidos / partido

Queremos que nuestro jugador reciba balones entre líneas debido a que esto es un parámetro de que se mueve bien buscando el espacio.

Vamos a ponderarlo con un 6.

7. Recuperaciones de balón/partido

Un mediocentro debe recuperar balones, sobre todo en un estilo de juego como el del Borussia de Dortmund con una presión fuerte en campo rival. Por ello creo que es un aspecto defensivo básico.

Vamos a ponderarlo con un 8.



8. Porcentaje de duelos aéreos ganados

El mediocentro que queremos también debe ganar un % alto de duelos aéreos, ya que son los encargados de bajar el balón al suelo cuando el portero rifa la pelota desde el saque inicial.

Vamos a ponderarlo con un 4.

6.7.2 Creación de un algoritmo de similitud para cada variable

La fórmula para calcular el porcentaje de similitud entre dos valores, x e y, es:

$$\text{Similitud } (x, y) = (\min (x, y) / \text{máx } (x, y)) * 100\%$$

En esta fórmula, se divide el valor mínimo entre el valor máximo y se multiplica por 100% para obtener el resultado como un porcentaje.

El resultado final indicará el grado de similitud entre los dos valores, siendo 100% una similitud total y 0% una similitud nula. Por lo tanto, cuanto más cercano sea el valor resultante a 100%, más similar serán los dos valores en cuestión.

Ejemplo: Queremos crear el % de similitud de npxG de cada jugador en relación con los que realiza Jude Bellingham, que es la variable la cual queremos comprar. Para ello, vamos a ver qué valor tienen Jude Bellingham, que es el cual vamos a querer comprar.

Realizamos la siguiente secuencia:

Seleccionamos la columna jugador -> Filtramos por nombre -> Buscamos Jude Bellingham

Vemos que tiene un valor de 0,21 npxG. Ahora ya sabiendo el valor que queremos comparar, vamos a crear una nueva columna y a aplicar la fórmula que podemos ver en la siguiente Ilustración 12.



Ilustración 12. % Similitud variable.

Vamos a realizar este proceso con el resto de los indicadores hasta tener el % de similitud de cada uno de los indicadores de cada jugador con los de Jude Bellingham.

6.7.3 Creación de algoritmo de similitud total

Una vez tenemos los % de similitud de todos los indicadores, lo que vamos a realizar es un % de similitud final ponderando cada uno de los mismos, con los valores que hemos decidido en los apartados anteriores.

Para ponderar una serie de valores y sacar la media ponderada, debemos asignar un peso o importancia relativa a cada uno de los valores (ya lo hemos hecho en el punto 6.6.1). Luego, multiplicamos cada valor por su respectivo peso y sumamos los resultados. Finalmente dividimos la suma obtenida entre la suma de pesos.

La fórmula para calcular la media ponderada es la siguiente:

$$\text{Media ponderada} = (w_1x_1 + w_2x_2 + \dots + w_nx_n) / (w_1 + w_2 + \dots + w_n)$$

donde:

x_1, x_2, \dots, x_n son los valores que se van a ponderar

w_1, w_2, \dots, w_n son los pesos o importancias que se les asigna a cada valor.

Esto lo vamos a poder ver de manera más clara en la siguiente Ilustración 13.

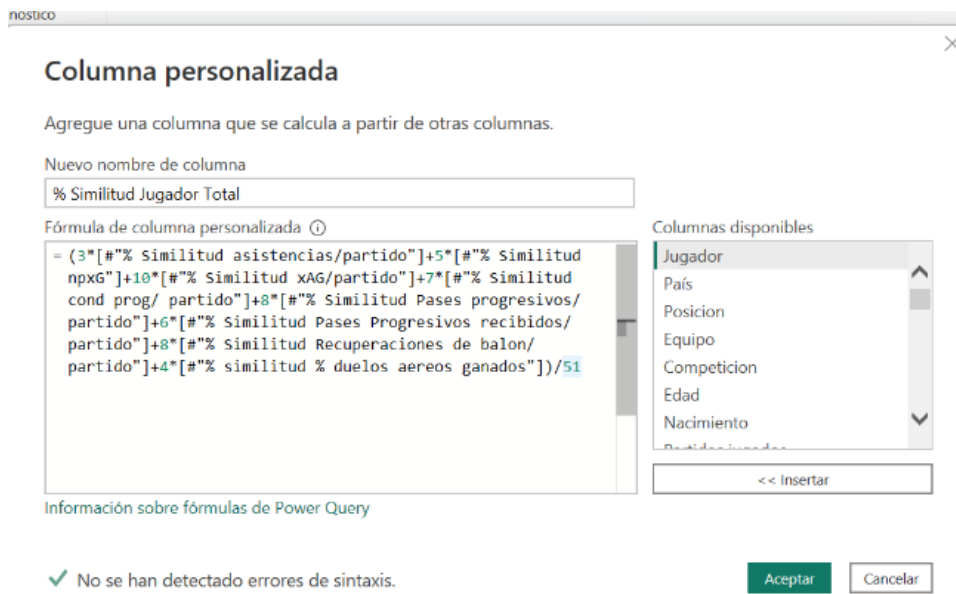


Ilustración 13. % de similitud total entre jugadores.

6.8 Visualización y cuadro de mandos

Una vez tenemos ya recogidos en Power Query todos los datos necesarios para obtener nuestras propias conclusiones lo que debemos hacer es crear nuestro propio cuadro de mandos para ver los resultados obtenidos y así mandar un informe a nuestro director deportivo.

6.8.1 Portada

Lo primero que vamos a hacer es realizar la portada de nuestro cuadro de mandos. Debe quedar claro cuál es el trabajo por realizar y quien es el que ha realizado el presente trabajo de detención del talento.

Ya que la herramienta va a ser para la dirección deportiva del Borussia de Dortmund todos los colores (amarillo y negro) y fotos de esta, van a ir relacionados. Para empezar, vamos a poner una imagen del estadio del Borussia de fondo. Para ello vamos a seguir el siguiente comando:

Pulsamos en el fondo del informe -> Dar formato a la página del informe -> Fondo del lienzo -> Imagen

La imagen la descargamos de internet y la guardamos en nuestro dispositivo.

Lo explicado anteriormente lo podemos observar en la siguiente Ilustración 14.

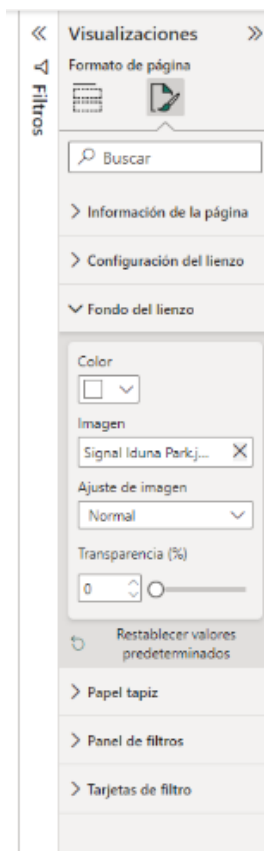


Ilustración 14. Fondo de pantalla Power Bi.

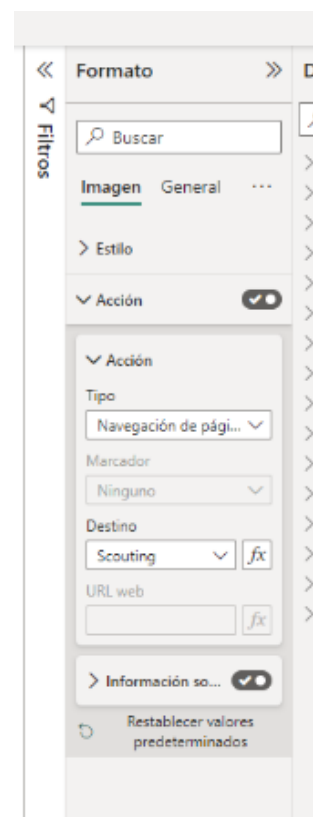


Ilustración 15. Botoneras.

Lo siguiente que debemos hacer es crear cuadros de texto para poder poner tanto el título como el autor. Para ello vamos a seguir el siguiente comando:

Insertar -> Cuadro de texto

Por último, lo que vamos a hacer va a ser una pequeña botonera para poder navegar mejor entre las diferentes páginas de la herramienta. Para ello vamos a descargar una serie de imágenes, las cuáles pondremos en pequeño y asignaremos a cada una un vínculo que nos lleve a cada una de las páginas.

Una vez tengamos insertadas las imágenes seguimos el siguiente comando:

Pulsamos en la imagen -> Dar formato -> Acción -> Destino

Lo podemos observar en la Ilustración 15.

Vamos a poner el escudo del club, para que nos lleve a la portada, la imagen de una búsqueda para que nos lleve a la pestaña de scouting de jugadores y la imagen de Jude Bellingham para que nos lleve a su algoritmo de similitud.

Una vez realizados todos estos pasos, el resultado de la portada va a ser el siguiente:
Ver Imagen 31.



Imagen 31. Portada principal Herramienta scouting Fuente: (Elaboración propia).

6.8.2 Algoritmo de similitud y scouting de jugadores

Las páginas de algoritmo de similitud y scouting de jugadores en Power BI se enfocan en analizar los datos y métricas más relevantes para evaluar el desempeño de los jugadores de fútbol, tales como valor de mercado, año de finalización de contrato, xAG, xG, asistencias, conducciones progresivas, pases progresivos, pases recibidos entre líneas, recuperaciones de balón o duelos aéreos entre otras. Estos son los KPI's que hemos definido en los apartados anteriores y que a través de ellos y sus respectivas ponderaciones, vamos a poder comparar a los diferentes jugadores.

Lo primero que vamos a hacer es definir un fondo. Como queremos que las visualizaciones sean legibles vamos a decidir un fondo de color amarillo, para que también como hemos indicado antes, sea un color que se identifique con el club.

Vamos a poner el escudo del club como logo en cada una de las páginas, así como las botoneras de salto de página que hemos puesto en la portada.

Una vez tenemos definida la base, vamos a establecer los parámetros de filtrado. Estos parámetros van a ser: Nombre del jugador, posición, edad, minutos jugados y % de similitud con Jude Bellingham. A partir de estos vamos a definir lo que queremos comparar en cada visualización.

Para ello vamos a seguir el siguiente comando:

Agregar datos a sus objetos visuales -> Segmentación de datos -> Seleccionamos el campo a filtrar

Ver Ilustración 16.

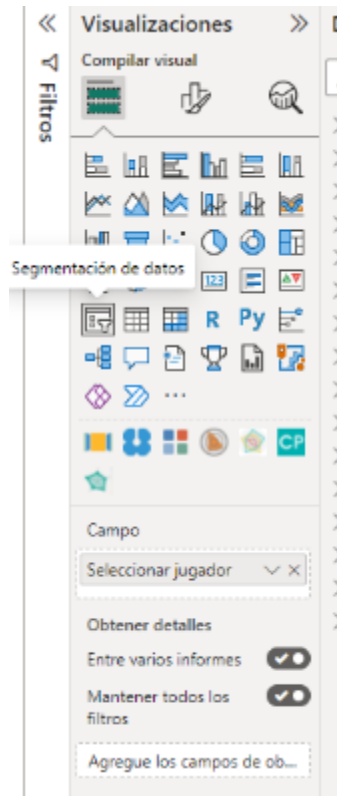


Ilustración 16. Segmentación de datos.

Para darle un formato acorde con la página, vamos a editar estas segmentaciones. Para ello vamos a seguir el siguiente comando:

Pulsamos en la segmentación a editar -> Formato Visual -> Objeto visual/general -> Texto/fondo/borde visual...

Ver Ilustración 17.



Ilustración 17. Dar formato a objeto visual.

Una vez tenemos el formato, vamos a definir estas segmentaciones para los valores alfabéticos y numéricos. Para los valores alfabéticos lo ideal es poner tanto un menú desplegable (en el cual incluya un control de búsqueda) como una selección múltiple, para poder seleccionar y comparar diferentes valores.

Lo vemos en la siguientes Imagen 32 , Imagen 33.

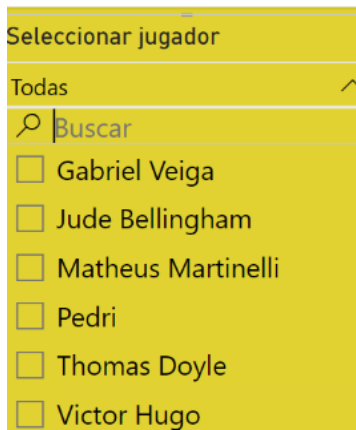


Imagen 32. Segmentación de búsqueda de jugadores.

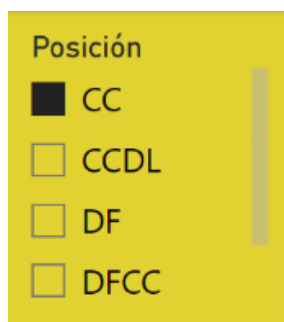


Imagen 33. Segmentación de posición de jugadores.

Para los valores numéricos lo más fácil es poner un control deslizante con el que poder acotar tanto la edad, como los minutos, como el % de similitud. Además, se crearán unas botoneras de forma automática a través de las cuales vamos a poder escribir el número que queramos. Lo vemos en las siguiente Imagen 34.

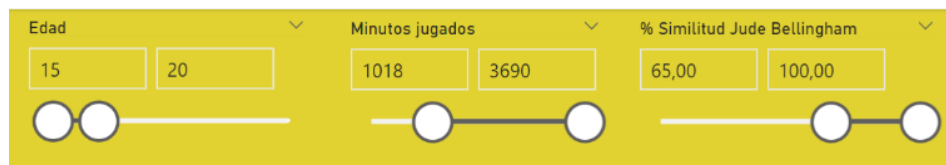


Imagen 34. Segmentación de edad, minutos jugados y % similitud.

En las anteriores imágenes, ya están reflejados los valores a acotar al realizar la selección. Queremos un jugador que juegue de centrocampista (CC), tenga menos de 20 años, ya que es la política de fichajes del club para posteriormente poder sacar una plusvalía, y que tenga un 65% de similitud al jugador que queremos reemplazar.

Una vez aplicados estos filtros, van a quedar 9 jugadores de los aproximadamente 5000 que tenía la base de datos inicial. Para que estos mismos filtros se apliquen en diferentes páginas en el informe, vamos a seguir la siguiente secuencia:

Ver -> Sincronización de segmentadores. Ver Ilustración 18.

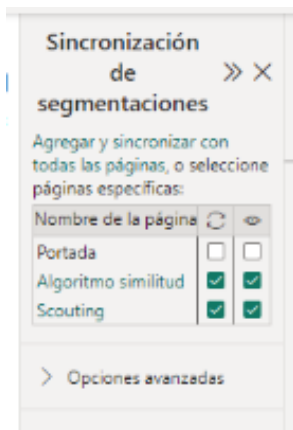


Ilustración 18. Sincronización de filtros.

A continuación, vamos a crear las diferentes visualizaciones para comparar los KPI's definidos a lo largo del proyecto.

Lo primero que vamos a realizar es una tabla.

En esta tabla vamos a ordenar de mayor a menor % de similitud con respecto al jugador que estamos comparando, ya que este % lo hemos creado nosotros a través de los valores que creíamos más importantes.

Después, para tener una visión más global de estos jugadores vamos a incluir, el equipo, la competición, la edad, y los diferentes KPI's que estamos comparando y vamos a comparar en las siguientes visualizaciones. En esta tabla también indicamos el valor de mercado (sacado de www.transfermarkt.es) y en el año en el que acaban contrato, ya que estos dos indicadores son bastante importantes a la hora de fichar a un jugador, como explicaremos más adelante.

Para realizar esta tabla, vamos a realizar el siguiente comando:

Visualizaciones -> Tabla -> Vamos añadiendo valores. Ver Ilustración 19

Para las siguientes visualizaciones ya no vamos a escribir el comando, ya que se hacen todas con el mismo.

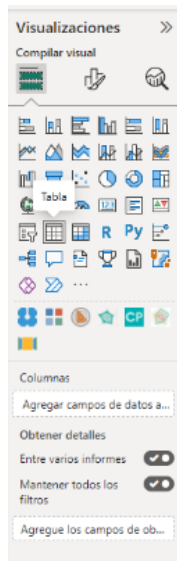


Ilustración 19. Agregamos tabla.

Dentro de esta tabla (Imagen 35), van a estar los nombres los cuales vamos a comparar a través de visualizaciones para después darle los que creemos que son los mejores sustitutos a nuestro director deportivo.

Jugador	% Similitud Jugador	Equipo	Competición	Edad	Minutos jugados	Valor de mercado (Mill.)	Finalización Contrato	xAG/90	Pases progresivos realizados/90	Recuperaciones/90	Conducciones progres
Jude Bellingham	100,00	Dortmund	Bundesliga	19	2255	120,00	2025	0,13	8,49	6,86	
Pedri	80,51	Barcelona	La Liga	20	1629	100,00	2026	0,18	8,12	6,30	
Matheus Martinelli	78,71	Fluminense	Liga Brasileña	20	1619	5,00	2025	0,12	6,61	5,33	
Gabriel Veiga	78,06	Celta Vigo	La Liga	20	1586	30,00	2026	0,13	4,43	5,30	
Victor Hugo	74,76	Flamengo	Liga Brasileña	17	1251	3,00	2027	0,15	4,24	4,39	
Johnny Cardoso	73,81	Internacional	Liga Brasileña	20	1410	3,50	2026	0,13	4,78	6,56	
Yunus Musah	69,06	Valencia	La Liga	20	1755	22,00	2026	0,08	4,46	4,38	
James Mcatee	67,44	Sheffield Utd	Championship	20	1491	6,00	2023	0,24	4,10	5,42	
Hayden Hackney	66,43	Middlesbrough	Championship	20	2416	2,00	2026	0,09	6,42	7,12	
Alex Scott	66,21	Bristol City	Championship	19	3098	14,00	2025	0,10	3,40	6,34	

Imagen 35. Tabla algoritmo jugadores. Fuente: (Elaboración propia)

Para observar de una forma más concisa este ranking de similitud, hemos realizado una gráfica de barras. En esta gráfica ordenamos los valores de mayor a menor % de similitud. Quiero destacar, que para comparar a Jude Bellingham, he dotado de color rojo a todas sus visualizaciones para que se vea de una forma más clara esa comparación.

Esto lo hemos realizado siguiendo el siguiente comando:

Pulsamos en la visualización -> Formato visual -> Colores

El resultado lo podemos observar en Imagen 36.

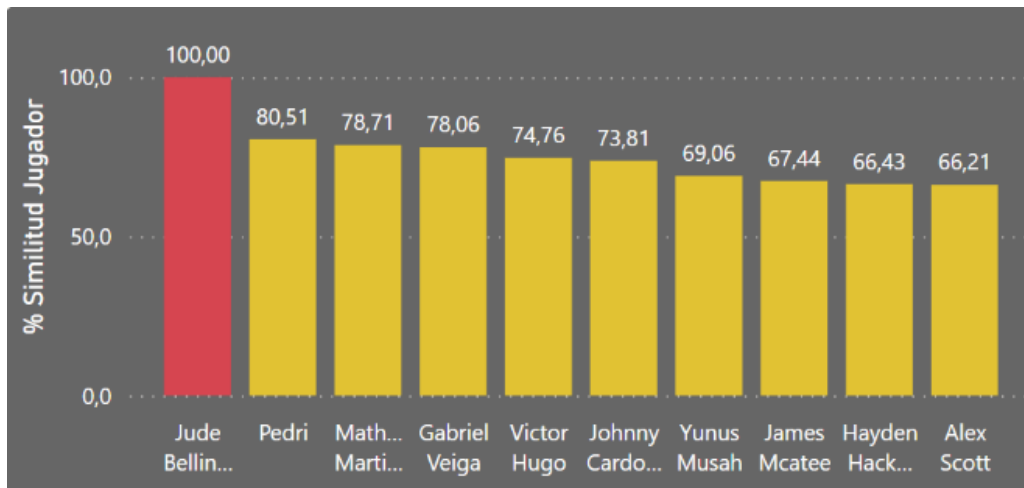


Imagen 36. Gráfico de barras. Fuente: (Elaboración propia)

Al tomar una decisión a la hora de fichar un jugador también debemos de darle información a nuestro director deportivo del valor de mercado de los jugadores seleccionados. Esta tarea es más propia de él que de los scouts, ya que el director deportivo se tiene que reunir con el director general y presidente o mandatario del club para ver el dinero que pueden gastar en fichajes y si ese fichaje es viable para que salgan las cuentas a final de año.

Nosotros servimos de ayuda para filtrar algunos de estos jugadores. Por ejemplo, en el actual caso, aunque Pedri sea el jugador con mayor % de similitud a Jude Bellingham, no vamos a darle un informe a nuestro director deportivo, ya que Pedri vale 110 millones de euros con una cláusula cercana a los 1000 millones después de su renovación la temporada pasada, luego es una operación inviable. Además, tenemos que informarnos de los años que restan de contrato a ese jugador, ya que cuánto más se aproxime la fecha, más barato será su costo para negociar con el club vendedor, ya que en vistas de que no va a renovar tienen urgencia por sacar dinero por ese jugador, y ante esa necesidad, es cuando nosotros aprovechamos la oportunidad.

Por ello este ranking lo podemos observar en el siguiente gráfico de líneas. (Imagen 37). Cuanto más abajo del gráfico crucen las 2 líneas, más barato nos va a salir ese jugador.

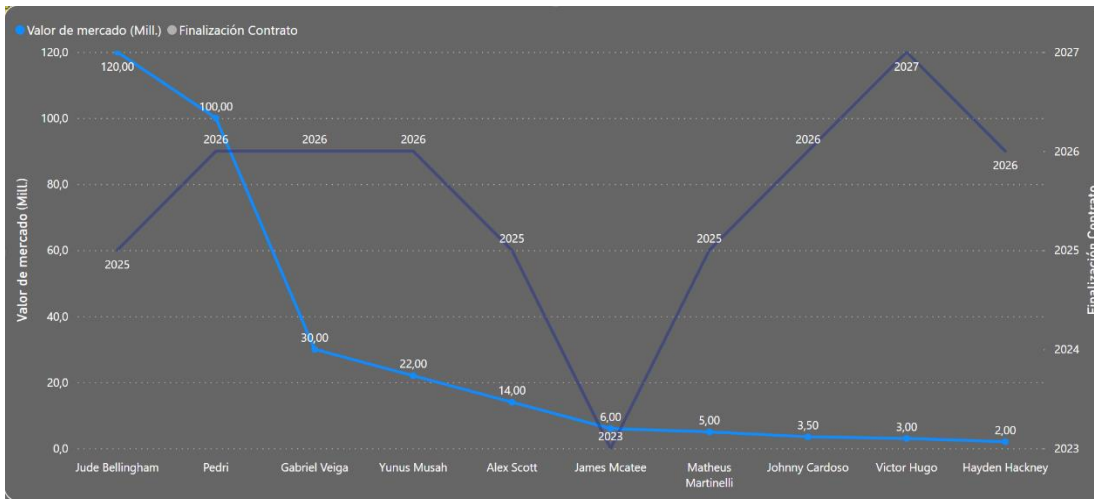


Imagen 37. Gráfico de líneas. Contrato + valor. Fuente: (Elaboración propia)

La página completa, nos queda de la siguiente forma:

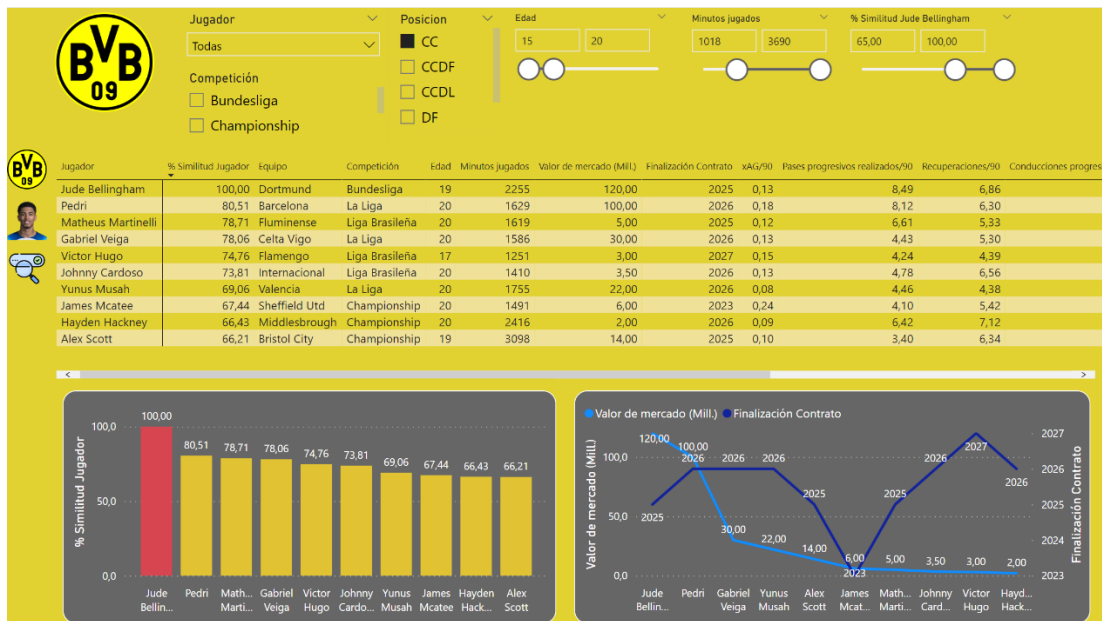


Imagen 38. Página completa algoritmo de similitud. Fuente: (Elaboración propia)

Una vez tenemos estos datos analizados, vemos que excepto Pedri, el resto de jugadores que nos han salido en lista, son a priori, económicamente viables, teniendo en cuenta además, que por la venta de Jude vamos a sacar una plusvalía. Por ello, ahora vamos a centrarnos en analizar los KPI's para seguir decantándonos por uno u otro jugador.

Para mí, la construcción de un gráfico de radar, es la forma más visual de relacionar todos los indicadores de rendimiento a la hora de comparar un jugador. Es útil para la comparación de 2 o como mucho 3 jugadores, ya que si no se van a juntar demasiados colores. Con este gráfico vamos a comparar a Matheus Martinelli con Gabriel Veiga, ya que son jugadores muy similares y son los 2 siguientes en la lista de similitud después de Pedri. Ver Imagen 39.

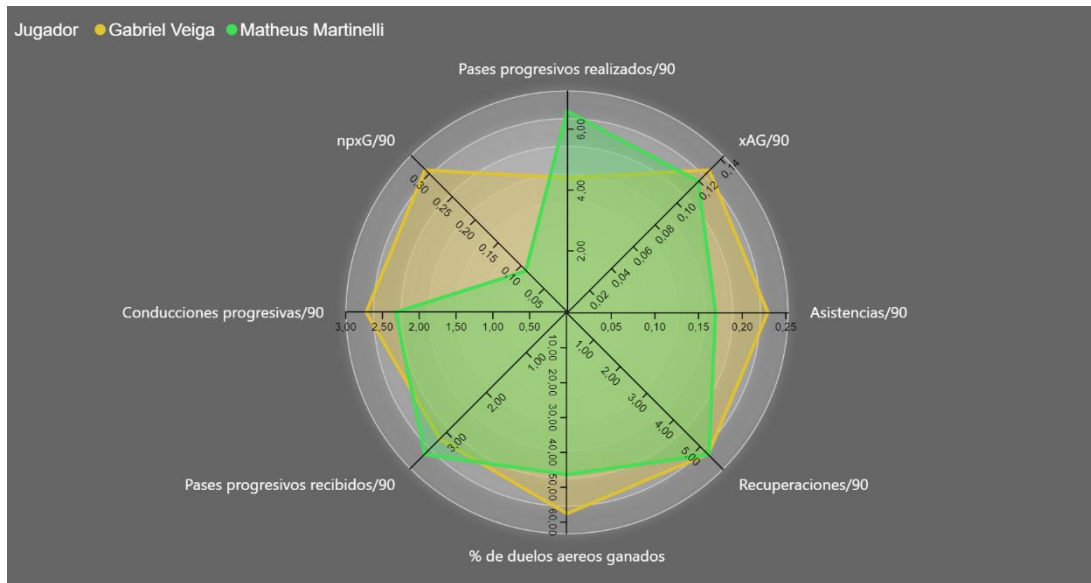


Imagen 39. Gráfico de radar. Comparación 2 jugadores. Fuente: (Elaboración propia)

Tras esta comparación, podemos ver que Gabri Veiga es mejor en todo excepto en pases progresivos realizados y pases progresivos recibidos. Con esto podemos llegar a la conclusión de que Gabri Veiga es un jugador más 'Box to Box' como se diría en Premier League, ya que no tiene tanto juego entre líneas y es un perfil más llegador. Mientras que Matheus Martinelli es un perfil con más toque de balón.

Además, debemos destacar el contexto de los dos jugadores. No es lo mismo los datos que nos da Gabri Veiga en una liga tan exigente como es la primera división española comparada con la primera división de Brasil. Además, la adaptación de un jugador brasileño a una liga europea es más costosa. Esto ya lo explicamos cuando hablamos de ligas de clase A y ligas de clase B.

Por tanto, a la hora de tomar la decisión, el mejor recambio para Jude Bellingham sería Gabriel Veiga, seguido de Matheus Martinelli.

Vamos a seguir analizando los otros posibles 3 nombres que van a aparecer en esa lista final a entregar a nuestro director deportivo.

Para ello deberíamos comparar a Victor Hugo y Johnny Cardoso, que son los 2 siguientes, pero al ser ambos de una liga menor, vamos a descartar uno de ellos.

Como ambos % de similitud son muy parecidos vamos a utilizar un Swarplot situacional o gráfico de dispersión para comparar las estadísticas tanto defensivas (% de duelos aéreos ganados y recuperaciones/90) como ofensivas como ofensivas de estos jugadores (xG y xAG), en comparación con un rango más amplio de jugadores de su misma posición y edad en otro contexto, como son las ligas europeas. Para ello mantendremos los filtros intactos, pero el % de similitud con Jude Bellingham lo quitamos, ya que no lo estamos comprando con él, y en el filtro de competiciones seleccionamos las 5 grandes ligas o ligas de clase A.

En la Imagen 40 podemos observar cómo tanto Johnny Cardoso (punto rosa) a nivel defensivo está por encima del promedio de jugadores de su edad y posición de las 5 grandes ligas en ambos indicadores, mientras que Victor Hugo (punto azul) está por debajo en ambos indicadores.

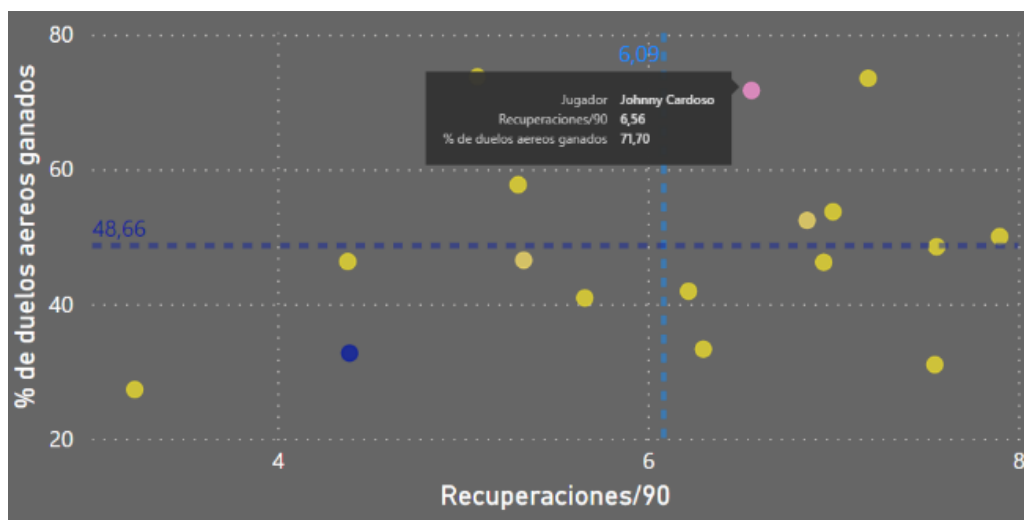


Imagen 40. Gráfico de dispersión datos defensivos. Fuente: (Elaboración propia)

En la comparación de las estadísticas más ofensivas, vemos que Victor Hugo está por encima de la media en ambos indicadores, al igual que Johnny Cardoso.

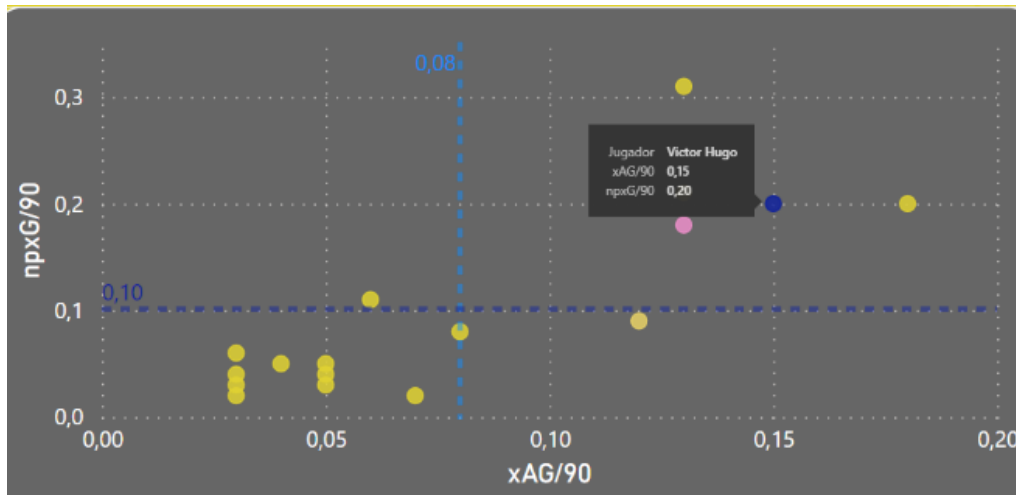


Imagen 41. Gráfico de dispersión estadísticas ofensivas. Fuente: (Elaboración propia)

A la hora de elegir uno de los dos, lo lógico sería quedarse con Johnny Cardoso, ya que tanto ofensivamente, como defensivamente está por encima de la media europea, mientras que Victor Hugo sólo lo está en la faceta ofensiva. Pero como en nuestro estilo de juego se le da más importancia a la fase ofensiva que defensiva y además Victor Hugo tiene tan sólo 17 años (y se va a poder sacar en un futuro una mayor plusvalía, además de un margen de mejora mayor al ser más joven) en comparación con los 20 de Johnny Cardoso, nos vamos a decantar por él como el tercero en la lista.

Ya sólo nos queda por dar 2 nombres más en nuestra lista de seleccionados.

De los siguientes nombres que nos quedan, al sólo estar Yunus Musah de ligas de clase A, vamos a quedarnos con él como cuarta opción. Además, aunque tienen % de similitud muy similares, es el que tiene uno mayor de los cuatro candidatos restantes con un 69,06%. También cabe destacar, que ante la mala temporada del Valencia CF, si finalmente descienden a segunda división, muchos jugadores querrán irse del equipo, y esta sería una buena oportunidad de mercado.

Ya sólo nos quedan 3 nombres por comparar: James Mcatee, Hayden Hackney y Alex Scott. Estos 3 jugadores, juegan en Championship, la segunda división del fútbol inglés. Aunque sea una liga de clase B, es una liga con un alto rango competitivo, así que la adaptación de estos a una liga mayor no creo que sea tan brusca. Además, tenemos el ejemplo de Jude Bellingham, que como comenté en apartados anteriores cuándo lo fichó el Dortmund provenía del Birmingham, de esta liga.

Para comparar a estos 3 jugadores he decidido utilizar un gráfico de barras apiladas horizontales (Imagen 42) en el cual se reflejen los aspectos ofensivos que rompen líneas defensivas rivales. Por ello he decidido comparar conducciones progresivas/90, pases progresivos realizados/90 y pases progresivos recibidos/90.

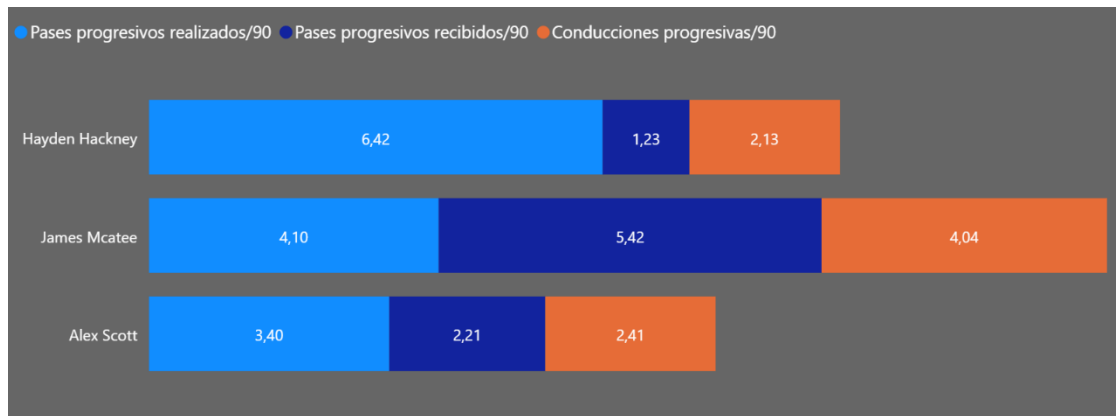


Imagen 42. Gráfico de barras apiladas horizontales. Fuente: (Elaboración propia)

Podemos comprobar que James Mcatee, es mejor que Alex Scott y Hayden Hackney, ya que su barra es la que más sobresale de las 3. Como estos indicadores son relativamente importantes para nuestro estilo de juego, vamos a descartar a Alex Scott ya que tiene un estilo más defensivo.

Para hacer una comparativa entre los 2 jugadores que nos quedan para elegir a uno, vamos a realizar una comparación con el gráfico de radar (Imagen 43), para ver los 5 indicadores restantes que no nos mostró el gráfico de barras y así por fin poder llegar a tomar la decisión.

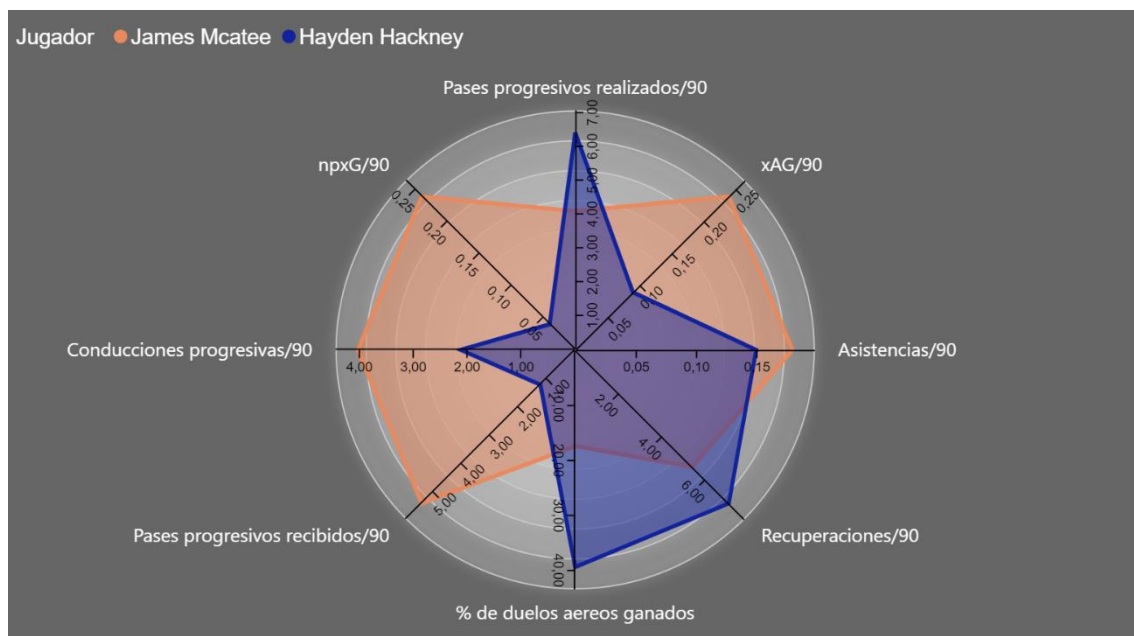


Imagen 43. Gráfico de radar. Comparativa 2 jugadores. Fuente: (Elaboración propia)

En esta comparativa podemos ver que James Mcatee es mejor que Hayden Hackney en 5 de los 8 indicadores de rendimiento propuestos. Además, en los atributos en los cuales es peor, es en los defensivos, que son a los que menos importancia les damos.

Por ello el 5 jugador que añadiremos a nuestra posible lista de sustitutos de Jude Bellingham será James Mcatee.

En la siguiente Imagen 44 tendremos una visión global de cómo ha quedado la página de scouting de nuestra herramienta de inteligencia de negocio deportivo.

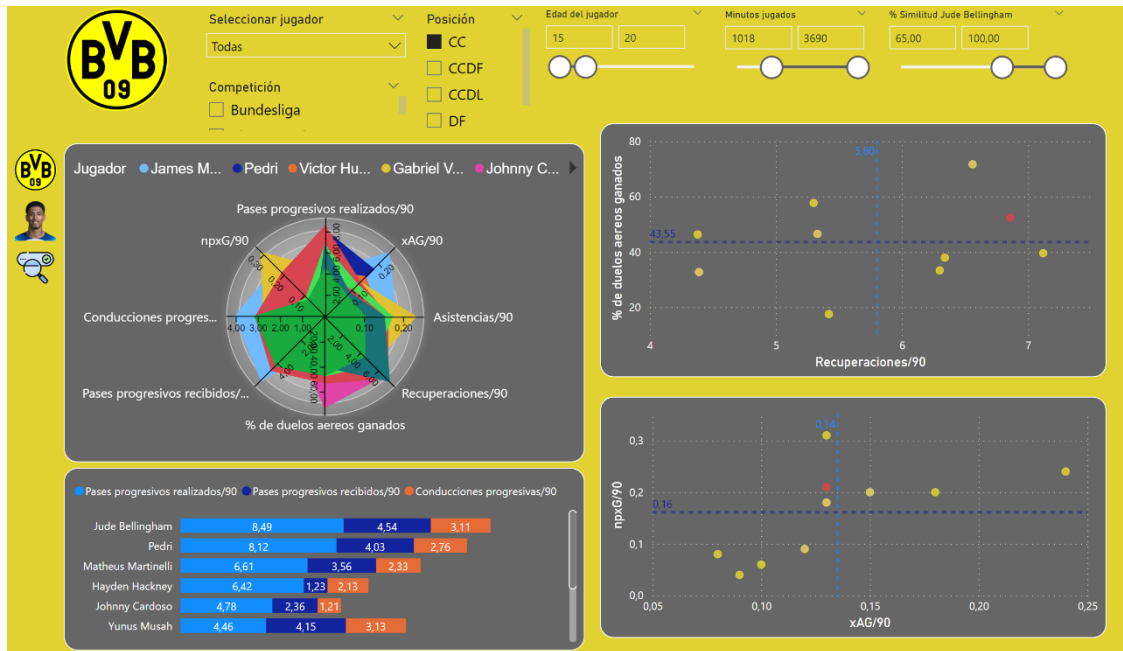


Imagen 44. Página completa scouting jugadores. Fuente: (Elaboración propia)



6.9 Informe de jugadores

Una vez obtenida la lista de posibles sustitutos, el siguiente paso es realizar un informe más completo de cada uno de ellos. Este informe, posteriormente se lo pasaremos a nuestro director deportivo, y él a partir de la lectura de este, tomará la decisión final.

Este tipo de informes se realizan para sintetizar todo lo que se ha estudiado a ese jugador (tanto objetivamente como subjetivamente) y que el director deportivo con un simple golpe de vista tenga toda esa información.

Vamos a realizar un ejemplo de informe visual del primer jugador de la lista, Gabriel Veiga.

1. Presentación del jugador



Imagen 45. Informe presentación Gabri Veiga.

Lo primero al realizar un informe es dar el nombre del jugador el cual vamos a estudiar y aportar una imagen de este. Además, en el informe debe aparecer el nombre del scout, (Sergio Machuca) por si hay más informes del mismo jugador realizado por diferentes scouts. Aparece el Escudo del equipo en el cual trabajamos y una imagen de fondo que representa la unidad del grupo. Ver Imagen 45.

2. Datos del jugador y visión subjetiva



Borussia Dortmund Player Report

GABRIEL VEIGA

Sergio Machuca Blázquez

Datos del jugador

Nombre completo: Gabriel Veiga Novas
Fecha de nacimiento: 27/05/2002
Lugar de nacimiento: Porriño (Pontevedra – España)
Edad: 20 años
Altura: 1,85m
Posición: Mediocentro/ Mediocentro ofensivo
Pie: Derecho
Agente: Gol internacional
Club Actual: Real Club Celta de Vigo
Contrato hasta: 30/06/2026
Valor actual de mercado: 30 millones

Visión subjetiva

Velocidad: 7
Técnica: 9
Capacidad física: 9
Visión: 8
Presión tras pérdida: 6
Juego aéreo: 7
Compañerismo (trabajo en equipo): 10
Repetición de esfuerzos a alta intensidad: 6
Reacción cuando sale desde banquillo: 9
Dinamismo: 7
Agresividad: 9
Finalización (tiro): 9
Movimientos sin balón: 9

Total/10: 8,07

Fuente de datos: Transfermarkt

Imagen 46. Datos del jugador y visión subjetiva.

En la siguiente página del informe tenemos que dar sus datos personales, así como el valor de mercado, y facilitar el agente o representante para que nuestro director deportivo se pueda poner en contacto con el mismo.

Realizamos una valoración subjetiva (escala del 1 (peor) – 10 (mejor)) de aquellos indicadores que consideramos importantes. Todos ellos salen de un visionado en directo de mínimo 3-5 partidos del jugador. Posteriormente se realiza una media de todos los indicadores. Ver Imagen 46.

3. Valoración objetiva (Big Data)

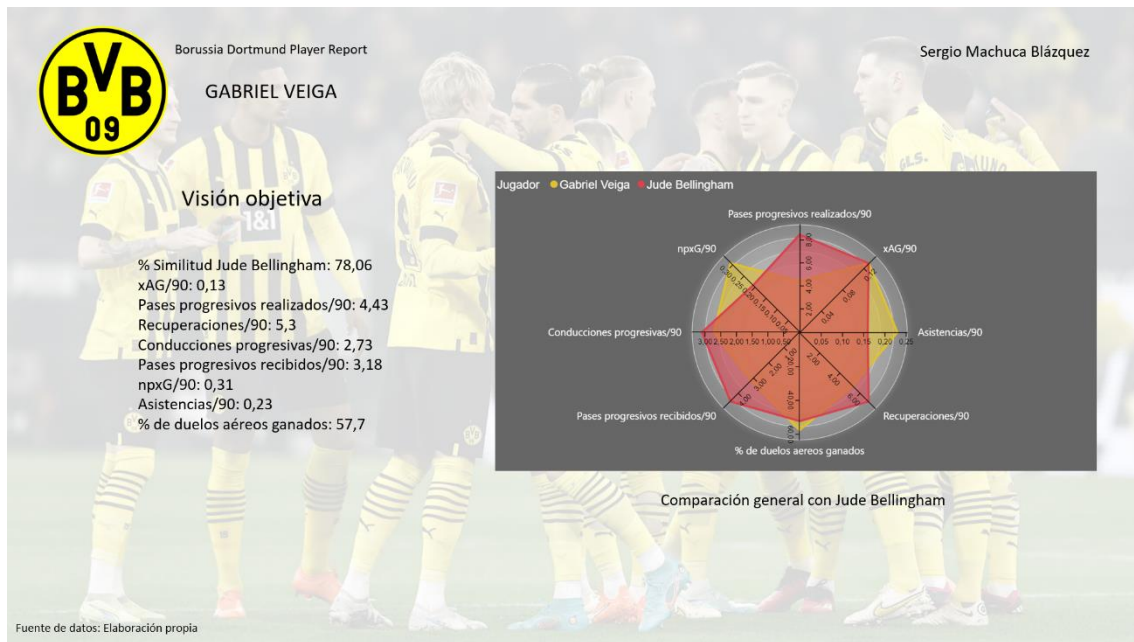


Imagen 47. Visión Objetiva.

Tras una visión más subjetiva del visionado en directo de partidos, sintetizamos los datos y conclusiones extraídas del estudio anteriormente realizado. Además, vamos a incluir el gráfico de radar, ya que aporta una comparación entre el jugador a sustituir y su sustituto de una directa.

Cabe destacar que nunca vamos a fichar a un jugador exclusivamente por los datos y tampoco vamos a fichar a un jugador sin contrastar lo visto en directo con los datos.

Por ello realizamos esa diferenciación entre las dos páginas en el informe. (Imagen 47)

4. Comentarios generales, decisión de fichar y vídeo



Borussia Dortmund Player Report

Sergio Machuca Blázquez

BVB
09

GABRIEL VEIGA

Comentarios generales

Más de 40 partidos a alto nivel en una liga tan exigente como la española.
Puede jugar tanto de mediocentro puro como de mediocampista con una vocación más ofensiva.
Sorprende la capacidad técnica a pesar de ser un jugador de 1,85 m.
Buen manejo con ambas piernas.
Es uno de esos chicos que cuando le ves jugar te entra por el ojo.
Gran capacidad para prodigarse en ataque llegando desde segunda línea lo que le hace marcar un gran número de goles.
Como opinión personal en un futuro se le podría probar como delantero centro debido a su envergadura y relación con el gol.
Se ofrece bien entre líneas y no le tiembla el balón.
Gran chico fuera del terreno de juego, opinión contrastada por varias personas cercanas a él.
No tendría problema en adaptación a la liga alemana, ya que Vigo precisamente no es una ciudad cálida.
Jugador salido de la cantera celeste, tan sólo 20 años y con gran proyección por delante.

DECISIÓN DE FICHAR: SI

Vídeo

<https://www.youtube.com/watch?v=h0E7vQK4qSc>
<https://www.youtube.com/watch?v=80B-GVTiNrs>

Fuente de datos: Elaboración propia

Imagen 48. Comentarios generales, decisión de fichar y vídeo.

Por último, vamos a realizar una serie de comentarios generales del jugador que no pueden estudiarse numéricamente.

Aportamos una serie de vídeos del jugador para que el director deportivo no tenga que perder el tiempo en buscar. Normalmente estos videos son realizados por el scout en función de lo que se quiera mostrar (diferentes indicadores estudiados). En el presente trabajo presento dos vídeos sacados de YouTube.

<https://www.youtube.com/watch?v=h0E7vQK4qSc>

<https://www.youtube.com/watch?v=80B-GVTiNrs>

Por último, decidimos si en base a lo que hemos estudiado ficharíamos o no ficharíamos al jugador de manera inmediata. Esto se realiza debido a que todos los informes quedan registrados en una base de datos propia del club, y el fichaje del jugador que estamos analizando puede postponerse en el tiempo. Ver Imagen 48.



7. RESULTADOS OBTENIDOS Y OPINIÓN PERSONAL

7.1 Resultados obtenidos

En este punto voy a valorar la consecución de los objetivos que fueron propuestos a principio del presente trabajo. Antes de centrarnos más a fondo en los objetivos secundarios, cabe destacar que el objetivo principal era el de realizar un cuadro de mando, fácil e intuitivo, que ayudase a un director deportivo de un club de fútbol en la toma de decisiones a la hora de fichar un jugador.

Creo que, como se puede comprobar en base a lo descrito a lo largo del apartado “Caso práctico” y a las capturas mostradas de la aplicación desarrollada, este objetivo se ha podido alcanzar de un modo satisfactorio.

Mediante la utilización de una serie de paneles desarrollados sobre la plataforma Power Bi y las visualizaciones interactivas en ellos implementadas, he conseguido acotar y seleccionar un rango de 5 jugadores candidatos de entre un número existente en la base de datos utilizada de aproximadamente 5000. Esta tarea de filtrado y selección previa, realizada en función de unos requisitos previamente marcados, debería permitir a nuestro director deportivo, no sólo el ahorrarse ese tedioso tiempo de búsqueda y procesado de una gran cantidad de datos, sino que, además, este pueda realizar su propio análisis de esos 5 candidatos basado en un estudio e interpretación más detallado de los datos ya procesados que son mostrados por las visualizaciones.

Centrándonos más en los objetivos secundarios, cuya consecución sucesiva nos han permitido llegar a alcanzar el principal, podemos destacar:

1. Conocer que es el mundo Big Data

Durante los primeros apartados del trabajo me he centrado en realizar un estudio del campo del análisis de Big Data desde un punto de vista global, no tan centrado en el mundo del fútbol.

La idea era mostrar la importancia que está tomando la recolección de las grandes cantidades de datos que se generan actualmente y cómo, la combinación de los métodos ya establecidos para el procesado de esas grandes cantidades de datos, conocidos como procesado Big Data, junto con el desarrollo emergente de técnicas como la Inteligencia Artificial, el Machine Learning, la Realidad Virtual... están facilitando sobremanera el análisis y la explotación de esos datos, hasta tal punto que se están convirtiendo en algo disruptivo en cualquier ámbito.

Entre ellos, el de Business Intelligence, es decir, en la tarea de transformar esta gran cantidad de datos en conocimiento útil para la toma de decisiones. Y el mundo del futbol no es ajeno a ello.



2. Conocer de primera mano cómo trabaja un club de fútbol a través de los datos

Una vez ya metidos en el mundo del deporte y más concretamente en el mundo del fútbol, explico cómo y de qué manera se están utilizando los datos en el mundo real. Ahora mismo no hay casi ningún club de fútbol a nivel profesional que no crea en el dato y en la importancia de este a la hora de tomar decisiones. Esto lo he reflejado con las diferentes imágenes del apartado “Big Data Deportivo” en las cuales se puede ver casos del mundo real en los que clubes de fútbol utilizan el dato. Además, hay cursos y másteres más específicos para formar a profesionales en el mundo del dato a nivel deportivo, ya que es el presente y futuro y cada vez se van a demandar más perfiles de este tipo.

3. Conocer las plataformas de obtención de estos datos.

En el apartado “Big Data en el fútbol” hablo de cómo se recogen estos datos dentro del mundo. Cabe destacar que hay empresas como Mediacoach, Huld, Statsbomb... que se encargan únicamente de la obtención de estos.

Además, cada club tiene sus propias formas de obtener los suyos a nivel interno como los microchips de los cuales también se ha hablado en este apartado.

Es importante conocer las fuentes de obtención de estos datos, ya que, por ejemplo, a lo largo del apartado “Caso Práctico” he utilizado datos Open Data (datos abiertos y gratuitos para todo el mundo) de páginas como www.fbref.com o www.transfermarkt.es.

4. Analizar una gran cantidad de datos a través de diferentes KPI's.

Dentro del apartado “Caso práctico” he detallado a través de las diferentes tablas, los diferentes indicadores a estudiar y porque elegimos estos y no otros para analizar el perfil de jugador que queremos.

Estudiar bien los indicadores que queremos medir es lo más complicado, ya que son los que van a reflejar los resultados que vamos a obtener y una mala elección de estos puede hacer que la decisión no sea la correcta.

Además, hay que comprobar cómo se reflejan estos diferentes indicadores a lo largo de todas las visualizaciones mostradas, y cómo a partir de ellos, es como se toman las decisiones.

5. Crear un algoritmo de similitud que compare a un jugador con el resto.

También, dentro del apartado “Caso práctico” he puesto en marcha la creación de un algoritmo de similitud, ya que para el caso que quería estudiar, como es la sustitución de un jugador, hay que fichar algo lo más parecido al que se va a ir de la entidad.



Por ello, a partir de diferentes parámetros y ponderando cada uno de los indicadores, se logra obtener un % de similitud en función de la importancia que se le da a cada uno de estos indicadores.

Este algoritmo de similitud es el punto clave a la hora de filtrar jugadores.

6. Conocer el programa Power Bi en su conjunto además de realizar un buen uso de este.

A lo largo del caso práctico he ido explicando paso por paso a través de diferentes ilustraciones cómo funciona este programa. He diferenciado tanto el programa Power Query, con el cual pude hacer el proceso de ETL, como el programa Power Bi Desktop, con el cual pude crear visualizaciones, las cuales también hemos ido mostrando a lo largo de este apartado.

7. Realizar un caso práctico dentro del mundo real.

Me he puesto en la situación de un caso real, ya que Jude Bellingham este año es muy probable que se marche a un club más grande que el Borussia Dortmund, y he analizado a partir de todo lo que sabía y el programa Power Bi cuales serían los 5 mejores sustitutos en cuanto a los datos.

Debido a mi estrecha relación con gente dentro del mundo del fútbol y el actual máster que estoy estudiando, he podido comprobar cómo trabajan los clubes a nivel profesional y he plasmado lo mejor que he podido todo lo aprendido, dentro de este apartado.

7.2 Opinión personal

Disfrutar haciendo lo que te gusta, intentar ser el mejor en ello y convertir esto en algo de tu día a día es mi objetivo.

Al realizar el presente trabajo me he dado cuenta de que cuándo eres apasionado de algo, el madrugar, estar delante de una pantalla horas y horas, o incluso salir fuera de tu zona de confort e irte a otro país no supone ningún esfuerzo.

La idea de relacionar el mundo del Big Data y la analítica avanzada con el fútbol, el cual es mi pasión, me venía rondando la cabeza tiempo atrás. He podido comprobar que los datos están inmersos en este mundo desde hace ya algún tiempo, pero que es algo que va a ir creciendo de manera exponencial. Con la investigación y divulgación durante estos meses de este mundo, he aprendido a contextualizar estos datos y a plasmar esta información dentro de un cuadro de trabajo. He podido aprender a cómo funciona el programa Power Bi, tanto Desktop como Query y como poder aplicar esto a un caso en el mundo real.

Además, me ha servido para saber que camino tomar en los próximos años y formarme para quien sabe, en un futuro, tener la oportunidad de poder trabajar en algo tan bonito como el fútbol.





REFERENCIAS BIBLIOGRÁFICAS

- ¿Cómo trabaja el Departamento de Análisis del Leeds United? (s. f.).
<https://www.nacsport.com/blog/es-es/Users/departamento-analisis-leeds-united>
- ¿Qué es Business Intelligence o inteligencia de negocios? (s. f.). Tableau.
<https://www.tableau.com/es-es/learn/articles/business-intelligence#:~:text=El%20Business%20Intelligence%2C%20BI%20o,decisiones%20basadas%20en%20los%20datos>
- Arjona, M., & Arjona, M. (2021). Sport Data Analytics: el dato que necesitas para ser Nº1 en el mundo deporte. *Canal Gestión Empresarial*.
<https://www.inesem.es/revistadigital/gestion-empresarial/sport-data-analytics/>
- Arroyo, A., & Arroyo, A. (2021, 30 junio). *Asistencias esperadas (xA): qué son y para qué sirven* - Driblab | Football powered by data' Driblab | Football powered by data'. <https://www.driblab.com/es/analisis-jugadores/asistencias-esperadas-xa-que-son-y-para-que-sirven/>
- Así trabaja el Área de Análisis del Sevilla FC (Parte 1). (s. f.).
<https://www.nacsport.com/blog/es-es/Interviews/descubriendo-el-area-de-analisis-del-sevilla-fc-parte-1>
- Bello, E. (2023, 15 febrero). *Big Data: qué es, para qué sirve y por qué es importante*. Thinking for Innovation. <https://www.iebschool.com/blog/valor-big-data/>
- Benítez, R., Escudero, G., & Kanaan, S. (2013). *Inteligencia artificial avanzada*.
Big Data: ¿En qué consiste? Su importancia, desafíos y gobernabilidad. (s. f.).
<https://www.powerdata.es/big-data>
- Bustios, M. N. (2022, 25 enero). *La importancia del Big data en el deporte*. ADN Deportivo. <https://adndeportivo.upc.edu.pe/2022/01/25/la-importancia-del-big-data-en-el-deporte/>
- Chen, H., Chiang, R. H. L., & Storey, V. C. (2012). Business Intelligence and Analytics: From Big Data to Big Impact. *Management Information Systems Quarterly*, 36(4), 1165. <https://doi.org/10.2307/41703503>
- Colaboración. (s. f.). *KPI's ¿Qué son, para qué sirven y por qué y cómo utilizarlos?*
<https://blog.es.logicalis.com/analytics/KPIs-qu%C3%A9-son-para-qu%C3%A9-sirven-y-por-qu%C3%A9-y-c%C3%B3mo-utilizarlos>
- Conesa Caralt, J., & Curto Díaz, J. (2012). Introducción al Business Intelligence. *En Editorial UOC (2a)*.



- De La C, J., De La C, J., & De La C, J. (2021, 23 marzo). Copernicus: el 'big data', un escudo contra el cambio climático. *El País*. <https://elpais.com/sociedad/2021-03-23/copernicus-el-big-data-un-escudo-contra-el-cambio-climatico.html>
- Dynamic. (2020, 25 septiembre). *Historia del Big Data. Descubre su origen y cómo se ha ido consolidando*. DYNAMIC. <https://www.dynamicgc.es/historia-del-big-data/>
- Estadísticas e Historia del Fútbol | FBref.com*. (s. f.). FBref.com. <https://fbref.com/es/>
- Fútbol Práctico. (2020, 12 mayo). *El Proceso de Scouting en un Equipo de Fútbol - Adrián Esparraga* [Vídeo]. YouTube. <https://www.youtube.com/watch?v=0lcjkevdmw8>
- Gallardo, C. (2023). Descanso y Fútbol: Beneficios para optimizar el rendimiento. *Blog de Fitness, Nutrición, Salud y Deporte | Blog HSN*. <https://www.hsnstore.com/blog/deportes/futbol/descanso/>
- Gómez, C. A. O. (2021). Beneficios y Usos del Big Data. *Revista Empresarial & Laboral*. <https://revistaempresarial.com/tecnologia/tendencias/beneficios-y-usos-del-big-data/#:~:text=%C2%ABActualmente%20los%20usos%20del%20Big,a%20partir%20de%20nuestras%20preferencias.%C2%BB>
- González Rodríguez, F. (2020). *Sandbox, aplicación del Big Data en el futuro del Fútbol*.
- Iruela, J., & Iruela, J. (2015). Qué es el Business Intelligence (BI) o inteligencia de negocios. *Canal Informática y TICS*. <https://www.inesem.es/revistadigital/informatica-y-tics/que-es-el-business-intelligence-bi-o-inteligencia-de-negocios/>
- Kimball, R., & Ross, M. (2010). *The Kimball Group Reader: Relentlessly Practical Tools for Data Warehousing and Business Intelligence*. John Wiley & Sons.
- Manuel. (2018, 6 octubre). *PLAYR: El primer sistema SmartCoach - Todo Sobre Camisetas*. Todo Sobre Camisetas. <https://todosobrecamisetas.com/2018/10/playr-el-sistema-smartcoach-que-se-toma-las-canchas/>
- Marca, R. (2022, 29 marzo). Víctor Orta: "Yo prefiero construir carreras que comprar carreras" *MARCA*. <https://www.marca.com/radio/2022/03/29/62430aa2e2704e66608b45a1.html>
- Rodríguez, F. J. G. (s. f.). La tendencia de los cuestionarios Wellness en el fútbol, ¿y un Wellness IA? *www.linkedin.com*. <https://www.linkedin.com/pulse/la-tendencia-de-los-cuestionarios-wellness-en-el-futbol-paco-gonzalez/?originalSubdomain=es>



Roy, P. P., Agarwal, A., Li, T., Reddy, P. K., & Kiran, R. U. (2023). *Big Data Analytics: 10th International Conference, BDA 2022, Hyderabad, India, December 19–22, 2022, Proceedings*. Springer Nature.

Smartlighting. (2023). *¿Utilizan las empresas españolas la inteligencia artificial y el big data? Nuevo informe ONTSI sobre el grado de aceptación de estas tecnologías*. <https://smart-lighting.es/utilizan-empresas-espanolas-inteligencia-artificial-big-data/>

Thorn, J., & Palmer, P. (2015). *The Hidden Game of Baseball*. University of Chicago Press.

Wikipedia. Borussia Dortmund (2023a) - Wikipedia, La Enciclopedia Libre, https://es.wikipedia.org/wiki/Borussia_Dortmund

Wikipedia. Edin Terzic (2023b) - Wikipedia, La Enciclopedia Libre, https://es.wikipedia.org/wiki/Edin_Terzi%C4%87

Wikipedia. Jude Bellingham (2023c) - Wikipedia, La Enciclopedia Libre, https://es.wikipedia.org/wiki/Jude_Bellingham

Wikipedia. Muthu Alagappan. (2023d) - Wikipedia, https://en.wikipedia.org/wiki/Muthu_Alagappan

Wikipedia. Sebastian Kehl (2023e) - Wikipedia, La Enciclopedia Libre, https://es.wikipedia.org/wiki/Sebastian_Kehl





ÍNDICE DE SIGLAS

BBDD: Bases de datos.

BI: Inteligencia de Negocio o Inteligencia Empresarial (Business Intelligence).

CC: Centrocampista

CD: Club Deportivo.

CF: Club de Fútbol.

CSV: Código Seguro de verificación (commaseparated values).

DF: Defensa

DL: Delantero

ETL: Extraer, transformar y cargar (Extract, Transform and Load).

FIFA: Federación Internacional de Fútbol Asociación (Fédération Internationale de Football Association).

IA: Inteligencia Artificial.

KPI: Indicador clave de rendimiento (key performance indicator).

LFP: Liga de Fútbol Profesional

NBA: Asociación Nacional de Baloncesto (National Basketball Association).

npxG: Goles esperados sin contra penaltis (non penalty expected goals)

PO: Portero

SDC: Sport Data Campus

SQL: Lenguaje de consulta estructurada (Structured Query Language).

TFG: Trabajo de fin de grado.

UD: Unión Deportiva.

UEFA: Unión de Federaciones Europeas de Fútbol (Union des Associations Européennes de Football).

xAG: Goles esperados con asistencia (expected assisted goals)

xG: Goles esperados (expected goals)

xPTS: Puntos esperados (expected points)





ÍNDICE DE TABLAS

Tabla 1. Valoración subjetiva. Fuente: (twitter (miguelangelcampos))	49
Tabla 2. Presupuesto	61
Tabla 3. Posición del jugador. Fuente: (Elaboración propia).....	70
Tabla 4. Edad Fuente: (Elaboración propia)	70
Tabla 5. Minutos jugados. Fuente: (Elaboración propia)	70
Tabla 6. Número de asistencias/partido Fuente: (Elaboración propia)	71
Tabla 7. npxG/ partido Fuente: (Elaboración propia)	72
Tabla 8. xAG/ partido Fuente: (Elaboración propia)	72
Tabla 9. Conducciones progresivas/partido. Fuente: (Elaboración propia).....	74
Tabla 10. Pases progresivos/partido. Fuente: (Elaboración propia)	74
Tabla 11. Pases progresivos recibidos/partido. Fuente: (Elaboración propia)	75
Tabla 12. Recuperaciones de balón/ partido. Fuente: (Elaboración propia)	75
Tabla 13. % de duelos aéreos ganados. Fuente: (Elaboración propia)	76





ÍNDICE DE IMÁGENES

Imagen 1. Distintas aplicaciones de la IA. Fuente: (Europa.rl, 2021).....	19
Imagen 2. Ventajas del Business Intelligence. Fuente: (Blog.Interius).....	23
Imagen 3. Ejemplo de cuadro de mandos interactivos. Fuente: (La magnética).....	24
Imagen 4. Película Moneyball (2011). Fuente: (FilmAffinity).....	28
Imagen 5. Aplicación HomeCourt. Fuente: (VivaBasket)	30
Imagen 6. Charlando con Víctor Orta en M2 Máster scouting. Fuente: (Elaboración propia)	31
Imagen 7. Analistas Manchester United Partido Europa League. Fuente: (VarskySports – Twitter)	33
Imagen 8. PLAYR (SmartCoach). Fuente: (todosobrecamisetas)	35
Imagen 9. Principales proveedores de datos técnico-tácticos. Fuente: (ObjetivoAnalista)	35
Imagen 10. Principales empresas que suministran dispositivos SmartCoach. Fuente: (ObjetivoAnalista).....	36
Imagen 11. Mapa de acción de Mbappe y Griezmann en un partido con su selección. Fuente: (Objetivo analista)	37
Imagen 12. Distribución de acciones Partido Real Madrid – Barcelona. Fuente: (MoneyBall Data).....	37
Imagen 13. Distribución de pases Minguenza. Fuente: (Driblab)	38
Imagen 14. Puntos esperados en el Partido Real Madrid – Barcelona. Fuente: (MoneyBall Data)	39
Imagen 15. % de presión realizada en cada zona del campo por el Cádiz CF. Fuente: (Marca)	39
Imagen 16. Mapa de disparos en visión frontal partido Real Madrid – Barcelona. Fuente: (MoneyBall Data).....	40
Imagen 17. Mapa de tiros visión panorámica. Joselu temporada 2020/2021. Fuente: Data Kicks.....	40
Imagen 18. Mapa de calor Espanyol (Partido Real Madrid – Espanyol) Fuente: (Marca)	41
Imagen 19. Diagramas de voronoi. Fuente: (Objetivo Analista)	41
Imagen 20. Carga por jugador. Fuente: (Objetivo Analista).....	42
Imagen 21. Cuestionario wellness. Fuente: (MoneyBall Data)	43
Imagen 22. Estudio de distancia recorrida a alta intensidad Coutinho en comparación con el resto de los jugadores. Fuente: (Objetivo analista).....	45
Imagen 23. Test de wellness control. Fuente: (Barcainnovationhub)	46
Imagen 24. Ojeador. Fuente: (Diario de Mallorca).....	47
Imagen 26. Herramienta de scouting deportivo para acciones ofensivas. Fuente (elaboración propia)	51
Imagen 26. Signal Iduna Park. Estadio Borussia Dortmund. Fuente: (Bundesliga).....	54
Imagen 27. Sebastian Kehl como director deportivo. Fuente: (Twitter)	55
Imagen 28. Sebastian Kehl como futbolista. Fuente: (Transfermarkt)	55
Imagen 29. Edin Terzic. Fuente: (Guetty Images)	57



Imagen 30. Jude Bellingham. Fuente: (BeSoccer)	59
Imagen 31. Portada principal Herramienta scouting Fuente: (Elaboración propia).	83
Imagen 32. Segmentación de búsqueda de jugadores.	86
Imagen 33. Segmentación de posición de jugadores.....	86
Imagen 34. Segmentación de edad, minutos jugados y % similitud.	86
Imagen 35. Tabla algoritmo jugadores. Fuente: (Elaboración propia).....	88
Imagen 36. Gráfico de barras. Fuente: (Elaboración propia)	89
Imagen 37. Gráfico de líneas. Contrato + valor. Fuente: (Elaboración propia).....	90
Imagen 38. Página completa algoritmo de similitud. Fuente: (Elaboración propia)	90
Imagen 39. Gráfico de radar. Comparación 2 jugadores. Fuente: (Elaboración propia)	91
Imagen 40. Gráfico de dispersión datos defensivos. Fuente: (Elaboración propia)	92
Imagen 41. Gráfico de dispersión estadísticas ofensivas. Fuente: (Elaboración propia)	93
Imagen 42. Gráfico de barras apiladas horizontales. Fuente: (Elaboración propia).....	94
Imagen 43. Gráfico de radar. Comparativa 2 jugadores. Fuente: (Elaboración propia)	94
Imagen 44. Página completa scouting jugadores. Fuente: (Elaboración propia)	95
Imagen 45. Informe presentación Gabri Veiga.	97
Imagen 46. Datos del jugador y visión subjetiva.....	98
Imagen 47. Visión Objetiva.....	99
Imagen 48. Comentarios generales, decisión de fichar y vídeo.....	100



ÍNDICE DE ILUSTRACIONES

Ilustración 1. Datos jugadores 5 grandes ligas.	62
Ilustración 2. Exportación de datos mediante CSV.	63
Ilustración 3. Power query.	64
Ilustración 4. Datos poco legibles.....	65
Ilustración 5. Distintos pasos aplicados.....	65
Ilustración 6. Datos legibles y ordenados.....	66
Ilustración 7. Cambio de título y eliminación de parámetros.....	67
Ilustración 8. Columna personalizada	68
Ilustración 9. Creación de nuevas métricas.....	73
Ilustración 10. Redondeo decimal.	73
Ilustración 11. Combinar consultas.	77
Ilustración 12. % Similitud variable.	80
Ilustración 13. % de similitud total entre jugadores.	81
Ilustración 14. Fondo de pantalla Power Bi.....	82
Ilustración 15. Botoneras.....	82
Ilustración 16. Segmentación de datos.	84
Ilustración 17. Dar formato a objeto visual.....	85
Ilustración 18. Sincronización de filtros.	87
Ilustración 19. Agregamos tabla.....	88

