

Título: Resultados y predicciones partidos FIFA últimas tres décadas.

Tabla de contenido

1. Introducción de datos	3
2. Descripción de la temática de datos	3
3. Definición Modelo relacional:	4
4. Diagrama relación de tablas	5
5. Listado de tablas	6
6. Listado tablas por columnas	7
7. Transformaciones realizadas	8
8. Análisis del tablero	11
9. Medidas calculadas	13
10. Conclusiones	14

1. Introducción de datos

La Copa Mundial de la FIFA es el torneo de fútbol más prestigioso del mundo. El campeonato se celebra cada cuatro años desde el inicio del torneo en 1930.

El formato de juego actual consta de una fase de clasificación, que tiene lugar durante los tres años anteriores, para determinar qué equipos pueden participar de dicho torneo donde, 32 equipos, compiten por el título en diferentes estadios del país anfitrión.

El pasado campeón fue Francia, que venció a Croacia en el torneo de 2018 en Rusia y en la versión de Qatar 2022, Argentina salió campeón en una final agónica contra Le blues el antiguo campeón Francia.

Este conjunto de datos ofrece una visión completa de todos los partidos de fútbol internacionales disputados desde los años 90, además, la fuerza de cada equipo se proporciona incorporando las clasificaciones reales de la FIFA.

2. Descripción de la temática de datos

Los datos utilizados provienen de una base de datos en Excel, que cuenta con casi 24000 entradas de todos los partidos FIFA jugados desde 1993, considerando, los partidos de cada mundial para tener en cuenta el clasificado del siguiente mundial por puntaje, por país anfitrión, clasificación directa o repechaje. la base fue dividida en varias hojas dependiendo del continente y el torneo a jugar. En África-CAF, Asia-AFC, Centroamérica-CONCACAF, Europa-UEFA, Latinoamérica-CONMEBOL y Oceanía-OFC.

Al contar con una base de datos con una línea de tiempo tan extensa y con una información tan detallada, se pueden realizar varios tipos de análisis, a continuación, veremos un breve análisis de cómo aplicar cada uno a nuestro tema:

- Descriptivo: ¿Qué pasó?, la base de datos cuenta con más de veintitrés mil partidos jugados desde el año 1993, nos cuenta con detalle que naciones se enfrentaron entre sí, dónde y cuándo.

- Diagnóstico: ¿Por qué pasó?, La base también cuenta con la cantidad de puntos a la fecha, así como el torneo por el cual se enfrentaron, ofreciendo información para crear un diagnóstico por continente.
- Predictivo: ¿qué pasará?, Gracias a la extensa base de datos se cuenta con un promedio de 113 partidos por selección, lo que puede ofrecer tendencias, medias, modas y medianas.
- Prescriptivo: ¿qué debería hacer?, de acuerdo con los resultados conseguidos en el pasado que resultaron en una victoria, se pueden realizar diversos escenarios que pueden ayudar a este resultado, (además de obtener la victoria en todos o la mayoría de los partidos).

3. Definición Modelo relacional:

Objetivo del proyecto:

Los miles de registros de partidos jugados durante los últimos 30 años en diversas competencias, pero todas con el mismo objetivo, la copa mundial de la FIFA se cuenta con la fecha, el lugar y el resultado de cada uno.

Con esta información, se puede analizar la tendencia en los resultados de cada juego, si se gana, pierde o empata de local y cómo influye la cantidad de goles anotados o recibidos.

Alcance y usuario:

Adicionalmente, se analizarán las tendencias de los países con más partidos ganados y diferencia en goles anotados, de esa forma se tiene una idea de cuanto alcance puede tener determinado equipo de acuerdo a las tendencias y proyecciones de los últimos encuentros jugados. Por contar con la información de todas las selecciones participantes hasta la fecha, cualquier persona interesada en el deporte puede utilizar la información, pero su enfoque va para los administradores de las casas de apuestas,

Esta aplicación de análisis es predictiva que los analistas utilizan a nivel operativo para

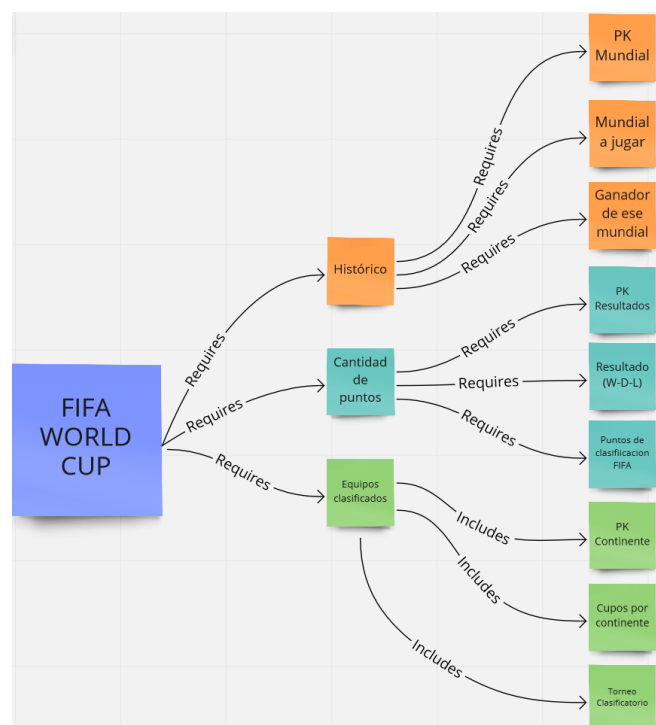
establecer los márgenes de ganancia y pérdida para cada apuesta, además de brindar información de observación a los entrenadores y deportistas.

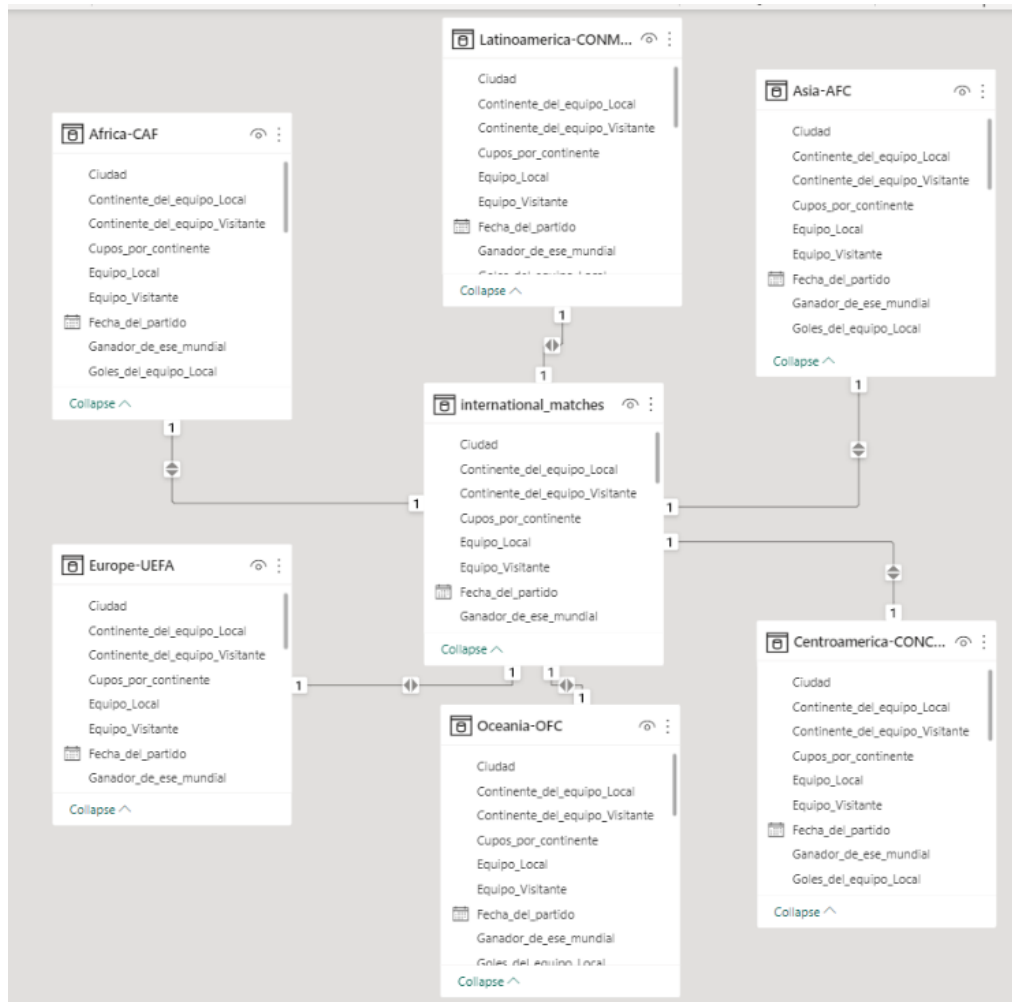
Entre los datos de la base de datos adicional a las tablas (que serán analizadas una a una más adelante), contamos con información precisa de ubicación, ciudad, país y continente. El torneo clasificatorio, el resultado del partido, la cantidad de goles recibidos y anotados, Y los puntos obtenidos por cada encuentro.

4. Diagrama relación de tablas

El diagrama de relación de tablas fue creado basado en la clasificación al mundial independiente de la cantidad de cupos o del repechaje. partiendo de la base del histórico para analizar tendencias, la cantidad de puntos para lograr avanzar y los equipos clasificados según el continente.

Cada tabla cuenta con una primary key, que puede definir las probabilidades de clasificación según los resultados, dentro de cada tabla, sí se tomó en cuenta los cupos y el torneo clasificatorio como un factor a considerar, lo cual se evidencia aún más en el modelado de las tablas de la base de datos.





Como fue mencionado anteriormente, la base de datos se dividió hoy en tablas acorde al continente y al torneo a jugar para la respectiva clasificación.

5. Listado de tablas

A continuación, vamos a visualizar las tablas encontradas en la base

Histórico

- PK_ID_MUNDIAL
- Mundial_a_jugar
- Ganador_de_ese_mundial

Cantidad de puntos

- PK_ID_Resultado_Local
- Resultados_del_equipo_Local
- Puntos_de_clasificacion_FIFA_del_equipo_Local

Equipos clasificados

- PK_ID_CONTINENTE_LOCAL
- Cupos_por_continente
- Torneo_Clasificatorio

6. Listado tablas por columnas

En cada tabla vemos el tipo de campo, su clave y su es Null.

TABLA 1 Histórico			
Campo	Tipo de campo	Tipo de clave	Null ?
PK_ID_MUNDIAL	int	PK	N
Mundial_a_jugar	int	X	N
Ganador_de_ese_mundial	int	X	N
PK_ID_Resultado_Local	int	FK	N
Resultados_del_equipo_Local	int	X	N
Puntos_de_clasificacion_FIFA_del_equipo_Local	int	X	N
PK_ID_CONTINENTE_LOCAL	int	FK	N
Cupos_por_continente	int	X	N
Torneo_Clasificatorio	int	X	N

Tabla 1: en esta primera tabla se obtiene la previsualización de los ganadores históricos de los mundiales jugados durante los últimos 30 años, entendiendo como factores determinantes los resultados de los partidos jugados durante la clasificación al torneo, así como los puntos de clasificación FIFA. De esta forma, obtendremos el histórico por cada país, encuentro y resultados, de acuerdo con los cupos por continente.

TABLA 2 Cantidad de puntos			
Campo	Tipo de campo	Tipo de clave	Null?
PK_ID_Resultado_Local	int	PK	N
Resultados_del_equipo_Local	int	X	N
Puntos_de_clasificacion_FIFA_del_equipo_Local	int	X	N

Tabla 2: La segunda tabla define la cantidad de puntos totales independiente de los triunfos pérdidas o derrotas que los equipos vivieron a lo largo del torneo y jornadas pre-clasificatorias, la finalidad de esta es comparar con los registros históricos y contra los datos del ganador del respectivo mundial.

TABLA 3 Equipos clasificados			
Campo	Tipo de campo	Tipo de clave	Null?
PK_ID_CONTINENTE_LOCAL	int	PK	N
Cupos_por_continente	int	X	N
Torneo_Clasificatorio	int	X	N

Tabla 3: la tercera tabla nos muestra a los equipos clasificados por cada continente de acuerdo con los puntos, partidos jugados, ganados, goles a favor y en contra, así como la diferencia de puntos. La finalidad de esta tabla mostrar la tendencia por cada continente y campeonato en donde Europa cuenta con más plazas en comparación al resto del mundo y las tendencias que estos manifiestan.

7. Transformaciones realizadas

La primera parte del tablero consta de un resumen general de la información contenida en la base de datos, cómo solamente se contaba con una tabla, esta se dividió por continente y su respectiva confederación teniendo en cuenta la cantidad de cupos con la que cada una cuenta por mundial. Esta información se basa en el actual sistema de

32 cupos en total, CAF en África que cuenta con 5 cupos, AFC en Asia con 4.5, la CONCACAF en Centroamérica que cuenta con 3.5, la CONMEBOL en Sudamérica con solamente 4.5 y la UEFA en Europa con 13 cupos.

Adicional a estos cambios, también se realizaron cálculos tanto en la base como en el tablero, para deducir, los puntos obtenidos con un sistema de ganar perder y empatar, proveyendo 3, 2 y un 1 respectivamente., así mismo se cuenta con sumatorias, totales generales categorizados por país continente y resultado.

Y finalmente, se realiza un análisis a detalle del actual campeón Argentina, con un histórico de los últimos 30 años, segmentación por fechas y otros que serán analizados más adelante.

- **Tabla International Matches**

Se genera consecutivo para cada mundial desde 1994, de igual forma se asigna número para cada resultado, sea ganar perder o empatar, (estos cálculos fueron aplicados a cada hoja divisoria por continente). para esto se crearon columnas calculadas para poder presentar la información relevante de cada país.

- **Tabla África-CAF**

La confederación africana de fútbol cuenta con 54 participantes y solamente 5 cupos, el país con más encuentros realizados es Suráfrica con 229 partidos, el equipo más ganador es Egipto con 163 partidos, Sudáfrica vuelve a aparecer como la selección más perdedora con 48 encuentros igualmente las que más ha empatado con 54. nos da un total de 5885 encuentros disputados por todas las elecciones durante los últimos 30 años.

- **Tabla Asia-AFC**

la confederación asiática de fútbol cuenta con 46 selecciones, y solamente 4.5 cupos para el torneo, con un total de 5302 partidos, la selección más ganadora es Arabia

Saudita con 164 partidos, con 64 partidos perdidos Catar y emiratos árabes ocupan el primer lugar con partidos más perdidos y con más empates se encuentra Japón con 63 encuentros que casualmente es la misma cantidad de partidos perdidos.

- **Tabla Centroamérica-CONCACAF**

La confederación de fútbol de la asociación del norte Centroamérica y el caribe cuenta con 34 participantes y 3.5 cupos para el mundial, con 2772 partidos en total desde 1994 el equipo con más derrotas es El Salvador con 65, con mayor cantidad de empates es México con 66 y al mismo tiempo la mayor cantidad de partidos jugados con 316. la selección con mayor cantidad de victorias es Estados Unidos con 194.

- **Tabla Europa-UEFA**

la unión de asociaciones de fútbol europeas cuenta con 55 países, pero un total de 13 cupos para el torneo cuenta con un total de 7593 partidos jugados durante los últimos 5 torneos en donde el equipo con más apariciones es Francia con 221, los equipos más ganadores con 145 encuentros lo ocupan Francia y España, los más perdedores son Luxemburgo y Malta con 95 encuentros y con 53 partidos la selección con más empates es Italia.

- **Tabla Latinoamérica-CONMEBOL**

La confederación sudamericana de fútbol cuenta con 4.5 cupos a la Copa del Mundo donde participan 10 selecciones, con un total de 1839 encuentros la selección con más cantidad de participaciones es Brasil con 233, siendo a su vez la más ganadora con 181 partidos, la mayor cantidad de empates tiene Bolivia con 39 encuentros y Venezuela con 46 partidos perdidos ocupa el último lugar.

- **Tabla Oceanía-OFC**

Finalmente, con 0.5 cupos la confederación de fútbol de Oceanía cuenta con 7 selecciones las cuales han tenido 530 encuentros, siendo Australia la mayor participante con 169 encuentros así mismo como la más ganadora con 112, la que más ha perdido

con 31 y la que más ha empatado con 26.

8. Análisis del tablero

A continuación, un análisis de los datos presentados en cada pestaña del proyecto en Power Bi, qué se puede encontrar en cada gráfico y cómo se realizó un análisis de la cual equipo campeón Argentina.

Pestaña 1: Portada

Resultados y predicciones partidos FIFA últimas 3 décadas, Argentina campeón.



A través de una base de datos que contiene partidos que datan desde 1994, de todos los encuentros FIFA oficiales realizados hasta la fecha (antes del mundial Catar 2022), se realizó un análisis predictivo que en su momento dio como resultado a un país latinoamericano campeón (Brasil).

En este documento se analizará las métricas de Argentina, durante los últimos 30 años.

Curso Data Analytics
Comisión 47645

Estudiante: Sebastian Vega Uribe



Pestaña 2: introducción de datos globales

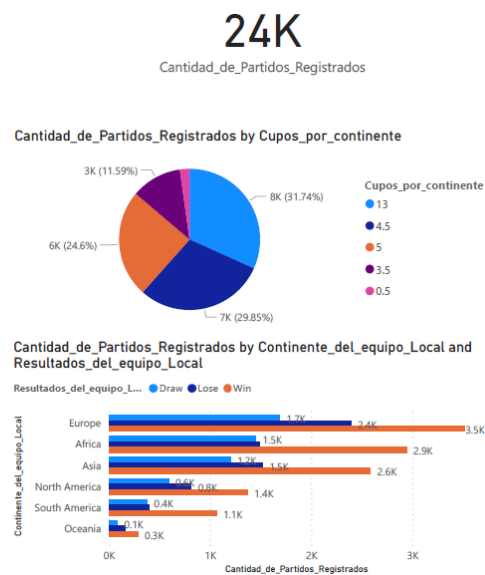
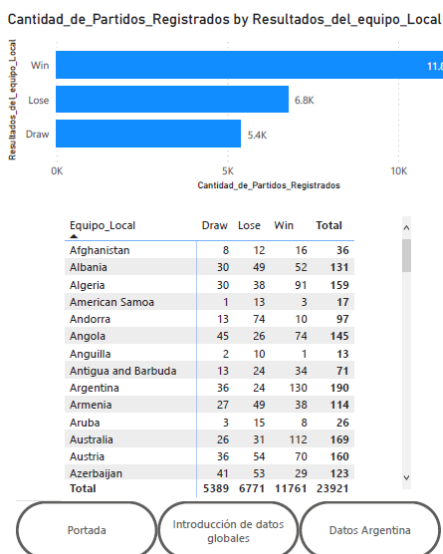
En la primera pestaña encontramos un resumen general donde nos menciona que durante las últimas 3 décadas se han jugado alrededor de 24.000 partidos (23.922 para ser exactos), dónde se ha marcado 64.048 goles de los cuales 38.494 son de local y 25.554 de visitante.

Presentación de datos globales



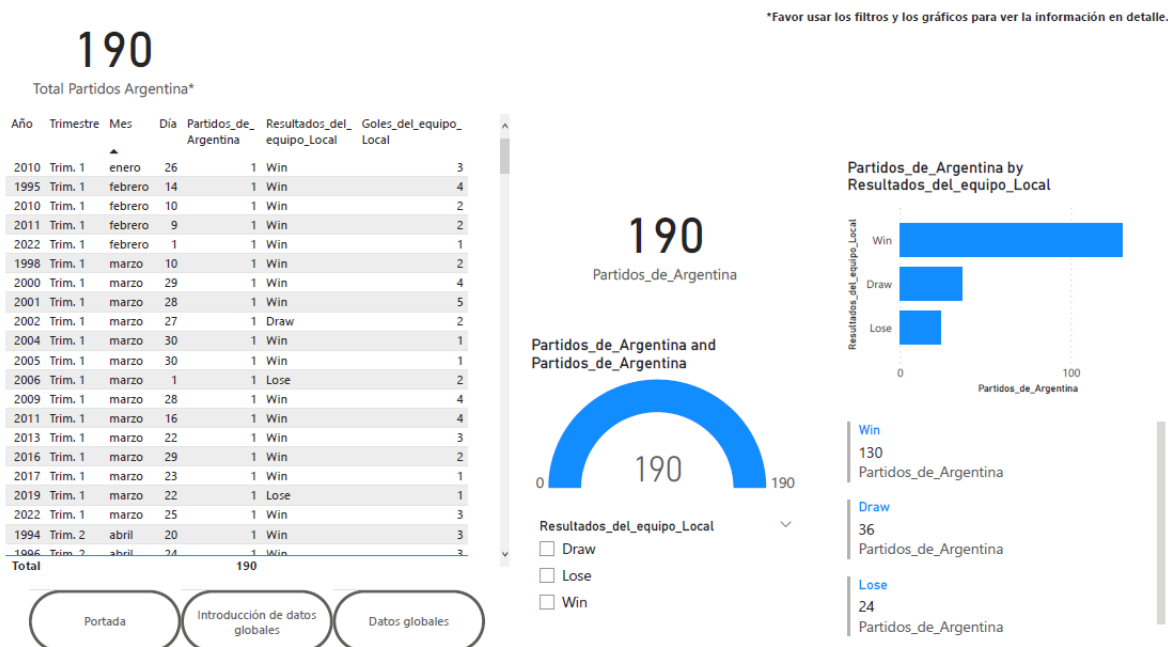
Pestaña 3: datos globales

En la siguiente solapa vemos con detalle cómo equipo se divide entre local y visitante, donde la información visualizada es del equipo local, y vemos cada categoría de ganar perder o empatar, a simple vista se nota la ventaja que tienen la selecciones jugando de local. Así mismo, vemos como hoy se distribuyen estas categorías por continente y como a lo largo de los años los cupos van jugando un porcentaje en cada copa.



Pestaña 4: datos Argentina

Finalmente, vemos en detalle la información de la selección campeona, desde 1994 Argentina ha jugado 190 partidos, de los cuales 130 resultaron en victoria. esta solapa cuenta con información interactiva entrante al hacer clic en alguno de los gráficos, hoy se puede ver la fecha la cantidad de goles la cantidad de encuentros y el resultado de dicho partido.



9. Medidas calculadas

Como se mencionó anteriormente, la mayor cantidad de datos calculados se encuentra en la solapa del análisis de Argentina, se cuenta con un campo calculado que suma la totalidad de los partidos jugados (190) desde 1994, como dato curioso de estos 190 ha ganado 130, empatado 36 y solamente perdidos 24.

Dentro de los campos calculados se encuentra una medida porcentual de los partidos ganados por la selección albiceleste, así como un campo específico solamente para este país, a continuación, las fórmulas:

Partidos_de_Argentina =

CALCULATE([Cantidad_de_Partidos_Registrados],international_matches[Equipo_Local] = "Argentina")

Medida_Porcentual =

CALCULATE([Cantidad_de_Partidos_Registrados],international_matches[Resultados_del_equipo_Local] = "Argentina")

Y finalmente se cuenta con una tabla de fechas en donde se puede visualizar año a año agrupado por semestre el día el resultado y la cantidad de goles realizados, lo cual muestra una tendencia de victoria es bastante alta durante los últimos 30 años.

10. Conclusiones

Realizar cualquier tipo de predicción en temas deportivos incluye factores aleatorios imposibles de pronosticar y considerando que el deporte más popular del planeta mueve cientos de masas y genera millones entre apuestas, mercancía, viajes y entradas, hoy es difícil ser racional frente a este tipo de datos, lo cual genera un grado más alto de dificultad a este tipo de proyectos inicialmente los resultados arrojaron como posible campeón a Brasil, siguiendo la tendencia que el último mundial jugado en territorio asiático fue ganado por éste, el segundo punto a considerar era que podía ser un país latino. De los 2 criterios solamente se logró uno, a pesar de contar con más de 24000 entradas de partidos desde 1994.

Aunque es posible realizar análisis de tendencias, y analizar la información de forma detallada los factores de azar influyen directamente en todos los modelos, El claro ejemplo de esto es Australia y Egipto. el primero cuenta con la mayor cantidad de partidos jugados y ganados de su torneo, solamente ha participado en 5 de 21 mundiales, donde nunca ha disputado un título ni llegado a una final, por su parte Egipto es la selección que más encuentros ha ganado, hoy solamente ha participado en 3 mundiales. Mientras que Marruecos con solamente 6 participaciones en el mundial Y

124 partidos ganados de 193 y su historia al convertirse en el primer país africano y árabe en llegar aún más semifinal de una Copa del Mundo con solamente un gol en contra y cuyas probabilidades iniciales como miembro del grupo F no eran las mejores ya que se encontraba con el seleccionado grandes como Croacia y Bélgica.

Por otro lado, los finalistas de la pasada copa mundial contaban con altas probabilidades ya que Francia como campeón del mundial de Rusia 2018 y Argentina eterna favorita disputaron un partido para el deleite del planeta, cuyas probabilidades de encuentro eran altas, pero favorecían al equipo europeo.