

Análisis descriptivo de pases realizados por jugadores de fútbol. Caso de estudio: Patriotas Boyacá

Descriptive Analysis of Passes Made by Soccer Players. Study Case: Patriotas Boyacá Team

*Ómar Alejandro Botía-Navas**

Fecha de recepción: 2 de octubre de 2018

Fecha de aprobación: 1 de noviembre de 2018

Resumen

Se ha hecho un análisis descriptivo sobre el comportamiento de los pases de los jugadores del equipo Patriotas Boyacá, en los partidos de la serie de treintidosavos de final de la Copa Sudamericana de Fútbol contra el Corinthians de Brasil, teniendo en cuenta el tiempo, el espacio y el partido dentro de la serie. Este trabajo se dividió en tres etapas: 1) definición de las características por recopilar, 2) extracción de los datos, y 3) análisis de los datos recopilados para describir lo sucedido. Este trabajo establece un formato para la extracción de datos generados a partir de partidos de fútbol y el análisis de los datos recopilados utilizando gráficas estadísticas y reportes interactivos de inteligencia de negocio. A partir de este análisis se entiende con mayor detalle el comportamiento de los jugadores sobre el espacio en un momento determinado de la serie.

Palabras clave: análisis descriptivo; fútbol; inteligencia de negocio.

Abstract

We have made a descriptive analysis over the behavior in the passes made by the players of the Patriotas Boyaca team, in the matches played against Corinthians on

* Universidad Pedagógica y Tecnológica de Colombia (Tunja-Boyacá, Colombia). omar.botia@uptc.edu.co.

the road of 64 by the Sudamericana Soccer Cup. This analysis was made taking into account variables like time, space and the match played in the series. This work was divided into three main steps, the first activity was the definition of the characteristic to gather, the second one was the data gathering, and finally we made the analysis over the data gathered to describe the behavior of the players in the matches. In this work we stipulate a format for the data gathering of soccer matches, we also used different statistical metrics and business intelligence tools to present and analyze easily the data. Throughout the data analysis we can understand closely the behavior of the players in a specific space and time on the stage.

Keywords: business intelligence; descriptive analysis; soccer.

Introducción

Entender cómo se comporta el rival y el equipo propio en las competiciones futbolísticas de alto rendimiento es una parte importante del juego. El conocimiento del equipo rival condiciona la forma de plantear un compromiso, para intentar disminuir sus fortalezas y explotar sus debilidades. Además, el conocimiento del equipo propio ayuda a planear sesiones específicas de entrenamiento y facilita la toma de decisiones en momentos determinantes de un partido.

Prueba de ello fue el cambio de arquero realizado al final del tiempo suplementario en los cuartos de final del Mundial de Brasil 2014 por parte del equipo de Holanda. Este cambio se debió a dos razones, una de conocimiento propio, la estatura del arquero que entraba era mayor que la del que salía, lo que le permitía abarcar una mayor área (BBC Mundo, 2014); la segunda, del conocimiento del rival, pues el arquero que entraba había estudiado hacia qué lado ejecutaban los cobradores del equipo rival, atajando así dos penaltis que ayudaron a su equipo a ganar la serie (Morenilla, 2014).

Previamente, en algunas investigaciones se analizó la red de pases realizados por los jugadores, dada la naturaleza de los datos con los que contaban no se hizo un análisis en el tiempo y el espacio (López & Touchette, 2013). El análisis espaciotemporal dividió el espacio de juego en diversas zonas en las que un jugador podía realizar o recibir un pase, además de llevar el registro del minuto en el cual se hacía la acción. Sin embargo, este trabajo presenta dos inconvenientes, las divisiones de la cancha son demasiado grandes, lo que hace perder precisión en el momento del análisis; además el estudio se llevó a cabo con llaves de eliminación directa y no con partidos de ida y vuelta.

En este análisis se dividió el campo de juego en 45 zonas, lo que aporta bastante detalle sobre el lugar originario de la acción y su destino; a su vez, se recopiló la información del minuto, el tiempo y la serie en la que sucedió el evento, para poder realizar

análisis en el tiempo de diferentes formas y poder comparar el comportamiento de los jugadores en los compromisos (Cotta, Mora, Merelo & Merelo-Molina, 2013).

Por medio de este trabajo se quiere demostrar la viabilidad del análisis en los pases de los jugadores de futbol a través de herramientas de analítica de datos que facilitan la visualización y el hallazgo de comportamientos en los jugadores, atendiendo diversas variables que condicionan el accionar del individuo.

Esta investigación aporta una metodología útil para el análisis del comportamiento en los pases de los jugadores de futbol. La metodología propuesta puede ser utilizada por cualquier club, en especial aquellos que no cuentan con los suficientes recursos para poder pagar un análisis a empresas especializadas.

Por medio de las visualizaciones y tableros de inteligencia de negocio generados a partir de la información recopilada se ha podido realizar el análisis descriptivo deseado, encontrando tendencias en el accionar del equipo en ambos partidos, así como diferencias en el juego al tener un resultado positivo en la serie o al sustituir un jugador.

El presente artículo se encuentra estructurado de la siguiente manera: se describe la metodología utilizada para llegar al análisis descriptivo de los datos. Se postula el mecanismo por utilizar para recopilar la información, las características para tener en cuenta y los dominios de las características. Se presenta el análisis descriptivo que se desprende de la visualización de la información a través de un análisis exploratorio de los datos y los tableros de inteligencia de negocio desarrollados; por último, se exponen las conclusiones de la investigación.

Metodología

Se define una metodología que es utilizada a lo largo del proceso, que facilita la organización y realización de las diferentes actividades. En primer lugar se definieron los datos por recopilar. El objetivo de esta etapa es determinar los atributos y sus dominios, buscando la facilidad para la recolección y que los valores recopilados fueran lo suficientemente detallados para poder realizar un buen análisis.

Al tener los datos recopilados se procede a la creación de visualizaciones y reportes de inteligencia de negocios que faciliten la abstracción de lo sucedido por medio de los productos generados, se realiza el análisis descriptivo de los datos, infiriendo las tendencias y las diferencias en el tiempo y en los diferentes momentos de la serie.

Los datos que se utilizaron para describir el comportamiento en este trabajo se recopilaron mediante la observación del partido y la anotación de forma manual en un

formato predefinido de las variables que se requieren para su posterior análisis. Se registraron los pases de jugadas con balón en movimiento y las acciones a balón parado (tiros de esquina, ejecución de faltas). En el caso del portero, se registraron los saques de arco y los pases que hizo con las manos. Entre los campos recolectados se encuentran los establecidos en la Tabla 1.

Tabla 1. Valores para los diferentes tipos de pase

Tipo de pase	Valor
Balón en movimiento	1
Saque de arco	2
Pase de manos	3
Balón parado	4

En una acción de pase debe haber un jugador que lo realiza (jugador origen) y un jugador que lo recibe (jugador destino); por tanto, se almacenó el número para el jugador de origen y el de destino, si el pase es interceptado o errado, el valor para el jugador de destino se registró como cero (0).

Respecto a las acciones de balón parado, se tuvo en cuenta si algún compañero realizó contacto con el balón, y de ser así, se marcó en el jugador de destino el número de su dorsal; por el contrario, si el balón salió de la cancha o fue interceptado por un jugador del equipo rival, el valor para el jugador de destino fue cero (0).

Se llevó un registro del minuto en el que sucedió la acción, para comprender el comportamiento de los jugadores a lo largo del tiempo; a su vez, se anotó si es del primer o segundo parcial del partido, evitando así la superposición de datos al medir las acciones en el complemento del primer tiempo (por ejemplo, el minuto 47 del primer tiempo y el 47 del segundo). De igual modo, al analizarse la serie completa se tuvo en cuenta si es del partido de ida o el de vuelta para diferenciar comportamientos tanto de local como de visitante.

Debido a que un mismo jugador puede actuar en diferentes zonas de la cancha en diferentes momentos, realizando pases para retener la posesión del balón o en busca del arco rival, se debe almacenar el lugar de donde procede y a dónde se dirige el pase.

Cotta *et al.* (2013) dividieron la cancha en nueve zonas (Figura 1); sin embargo, como las zonas establecidas son demasiado grandes se perdió precisión en los datos, ya que puede suceder que un volante central tienda a generar pases recostado a una banda, sin tener que estar completamente junto a la línea lateral, ya sea porque él recupera el balón en esa zona, se genera algún tipo de rebote, u otro compañero recupera el balón y lo utilice a él como salida inmediata; de igual forma, no es lo mismo que un lateral genere un pase apenas pasando la mitad de la cancha a que llegue a línea de fondo y

tire un centro, para este tipo de casos se propone utilizar una división de la cancha con 45 zonas demarcadas como se muestra en la Figura 2.

El cruce de las limitaciones de largo y ancho de la cancha se pueden ver como una matriz, cada celda refleja una posición en el terreno, ya sea de origen o destino de un pase. Su notación será primero el valor del largo, seguido del valor del ancho, contando el ancho de derecha a izquierda y el largo de arriba abajo, siendo así que el valor 11 hará referencia banderín de córner derecho del arquero, la posición 53 hará referencia al centro del campo, el valor 95 será el banderín de córner izquierdo.

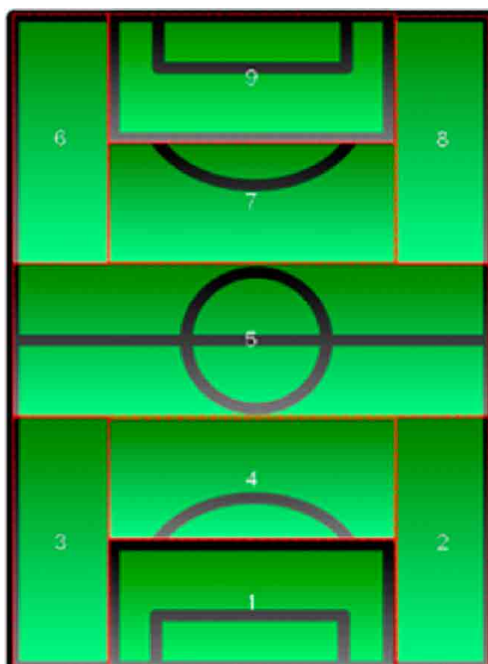


Figura 1. División del campo.

Fuente: Cotta, Mora, Merelo y Merello-Molina (2013).

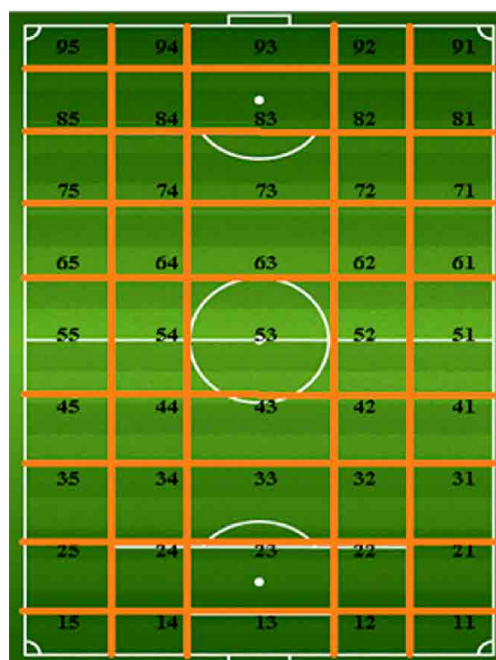


Figura 2. División de la cancha propuesta.

Teniendo en cuenta las anteriores precisiones, el archivo generado tiene los atributos de: *partido*, *tiempo*, *minuto*, *tipo de pase*, *jugador origen*, *zona origen*, *jugador destino*, y *zona destino*.

Una vez finalizada la recolección de los datos se procede a su análisis, utilizando herramientas como MS Power BI y R a través del IDE RStudio. Estas herramientas son altamente utilizadas en el análisis de datos además de ser *opensource*.

Resultados y discusión

Los archivos obtenidos fueron almacenados en variables denominadas *data frames* en R, estos *data frames* se combinan para utilizar un *data frame* con la información consolidada. La cantidad de observaciones está distribuida como se describe en la Tabla 2.

Tabla 2. Número de registros en los *data frames* en R

Primer partido	Segundo partido	Consolidado
390	454	844

A los ocho atributos se adicionó la columna minutos.bucket, la cual agrupa por un rango de minutos un grupo de acciones, los rangos son: 0-15 min, 15-30 min, 30-terminación del tiempo. Estos grupos tienen límite inferior excluyente y límite superior incluyente. Para el análisis de los datos se transformaron a factores todas las columnas, a excepción de la columna “Minutos”.

A pesar de que el segundo partido fue jugado en condición de visitante, el número de pases realizados fue mayor, esto indica que la continuidad del juego fue más alta. Este es un comportamiento esperado, ya que el equipo Corinthians juega normalmente sobre una altitud relativa baja al nivel del mar, así que al jugar en una ciudad como Tunja intenta bajar el ritmo de juego.

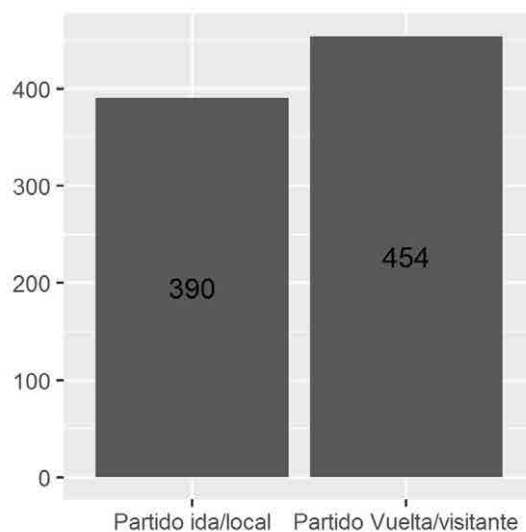


Figura 3. Número de pases por partido.

En el primer encuentro, Patriotas Boyacá hizo más pases en el primer tiempo, sin una diferencia excesiva en comparación con el segundo. En el partido de vuelta

se encuentra una diferencia notoria entre los pases realizados en cada uno de los tiempos, con 276 pases en el segundo tiempo y 178 pases ejecutados en el primer tiempo. Esta conducta está condicionada por la presión que impuso Corinthians en la primera media hora de partido, hasta anotar el gol de la ventaja, desde ese momento Corinthians entregó el protagonismo del juego a Patriotas esperando una acción para liquidar la serie, acción que llegó al minuto 46 del segundo tiempo.

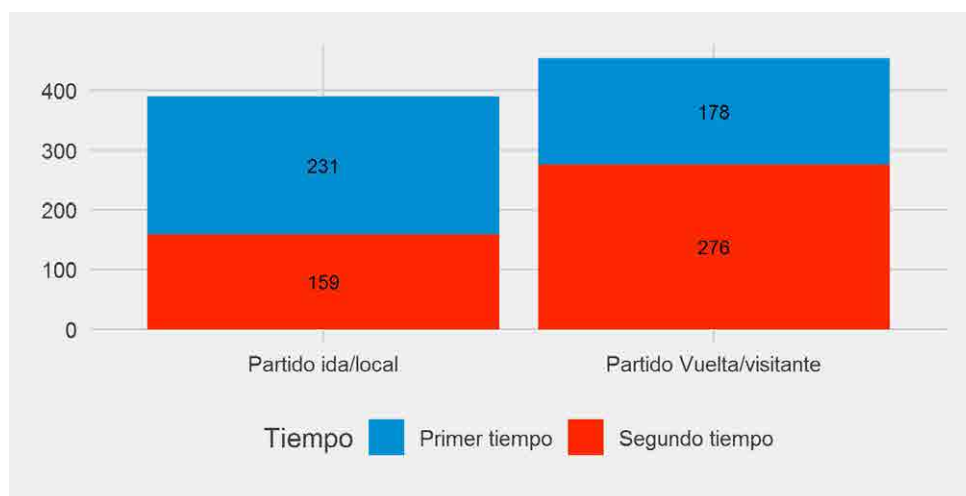


Figura 4. Número de pases por tiempo y partido.

En el segundo tiempo del partido de ida, el equipo Patriotas Boyacá optó por entregar la posesión del balón al rival e intentar aprovechar el espacio generado por el equipo rival, ya que en este momento de la serie se encontraba por debajo en el marcador. Este comportamiento se ve modificado en el segundo tiempo, cuando Corinthians le entrega la posesión del balón al equipo Patriotas Boyacá que tiene que buscar un gol para empatar la serie.

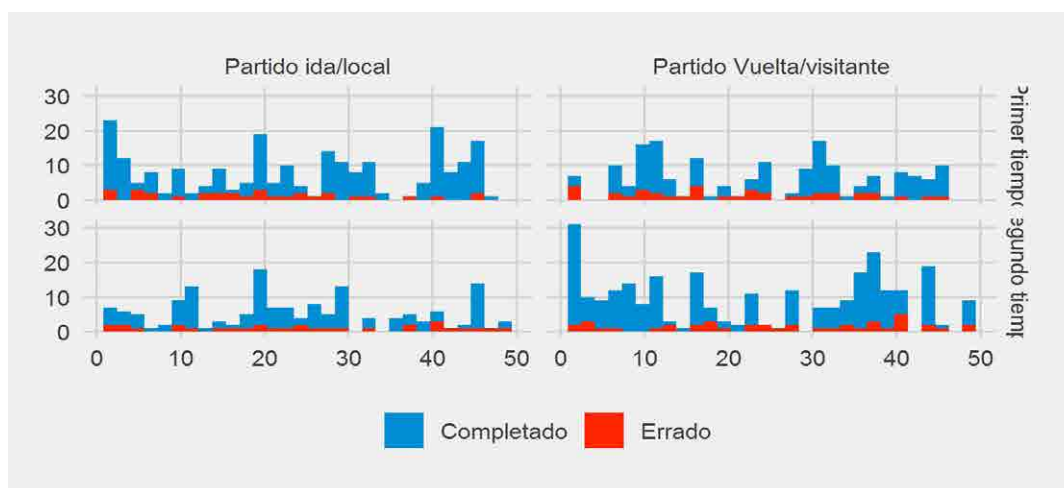


Figura 5. Número de pases por minutos según el tiempo y partido.

En el partido de vuelta se erraron 17 pases más que en el partido de ida, donde la mayor tasa de pases fallados se encuentra entre los 15-30 minutos del segundo tiempo del partido de ida; sin embargo, el periodo en el que más pases se erraron fue del minuto 30-50 del segundo tiempo en el partido de vuelta, con un total de 19. La menor tasa de pases errados se encuentra entre los 30-45 minutos del primer tiempo del partido de ida con tan solo seis pases fallados.

Para el primer partido se observa que hubo una fuerte relación entre Ómar Vázquez (10) y el jugador Larry Vázquez (5) con más de 15 pases; se evidencia una conexión interesante desde Rafael Robayo (8) a Carlos Mosquera (27) con 12 pases exitosos. El jugador que más pases erró fue Carlos Mosquera (27), con 12 fallos.

En el segundo encuentro los rangos de pases fueron más amplios entre los jugadores, se observa que entre Rafael Robayo (8) y Larry Vázquez (5) hubo más de 25 pases; así mismo, el número de pases errados aumentó considerablemente en este encuentro, ya que hubo más jugadores que erraron más de 10 pases.

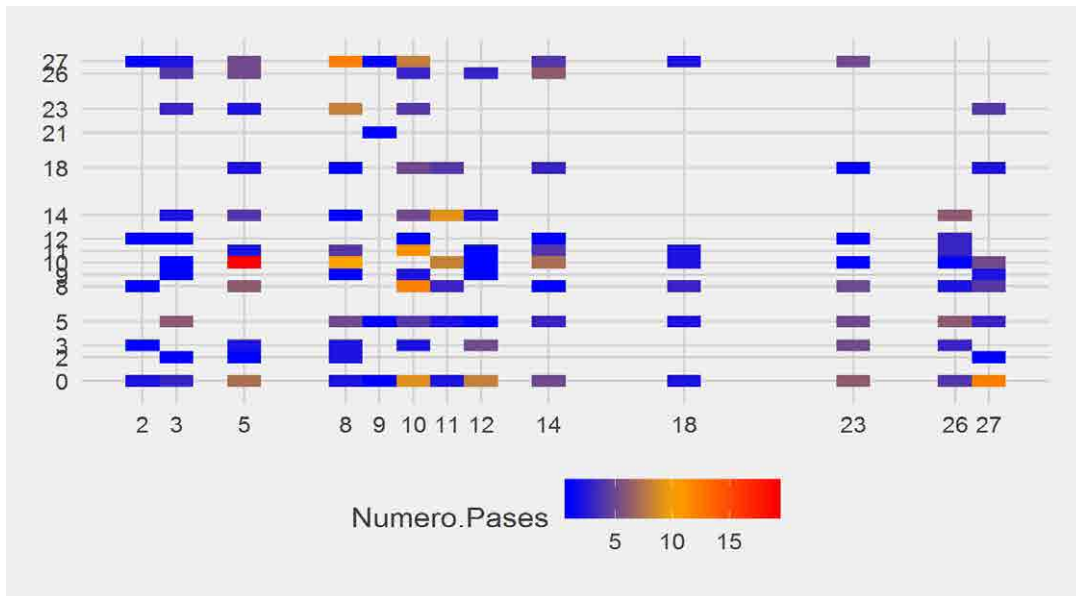


Figura 6. Relación de pases por jugador de origen y destino partido de ida.

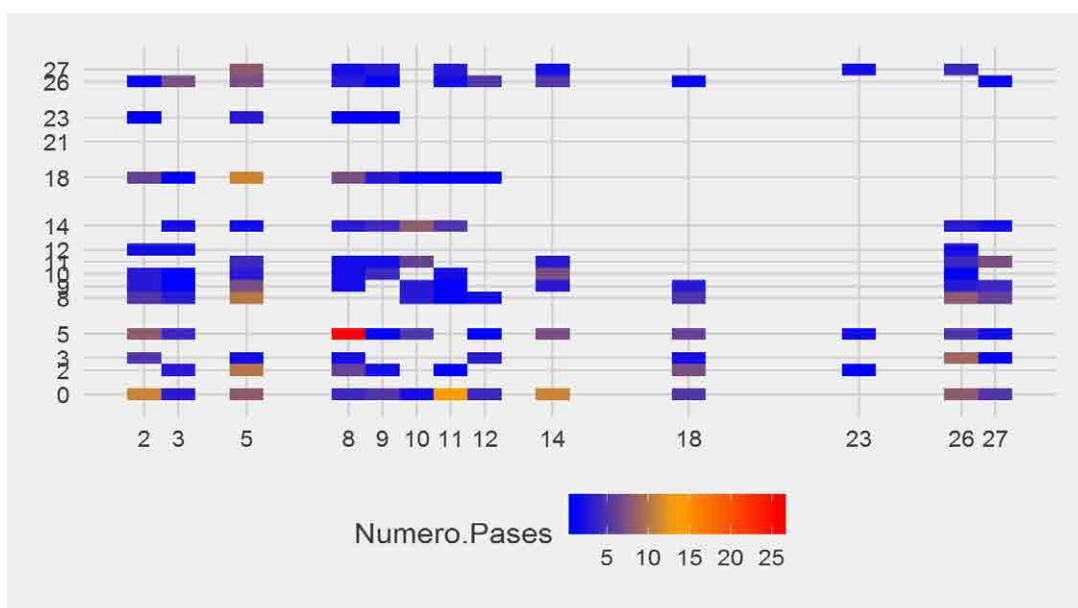


Figura 7. Relación de pases por jugador de origen y destino partido de vuelta.

Adicionalmente, del comportamiento de los jugadores según su rol o posición en el campo se logró determinar lo siguiente:

Análisis partido de ida

Centrales

Los pases recibidos y realizados por la pareja de centrales para el primer compromiso fueron equitativos. Para el segundo partido, Óscar Cabezas (26) tuvo un mayor contacto con el juego, recibiendo más de 9 veces el balón y ejecutando 20 pases más que su compañero de saga, este crecimiento en el número de pases se evidencia en el segundo tiempo.

Mediocampistas

- **Larry Vásquez (5):** realizó más de 20 pases en el segundo partido para el segundo tiempo de este compromiso, dada la salida en el entretiempo de Ómar Vásquez, Larry ejecutó más del doble de pases en comparación con el primer tiempo (51 vs. 24) y fue más buscado como opción de salida por sus compañeros, recibiendo 44 pases en el segundo contra los 23 recibidos en el primero. Con una mayor participación en el juego, Larry erró más pases (2 en el primero, 6 en el segundo)

- **Rafael Robayo (8):** al no encontrarse Ómar Vásquez en el terreno de juego para el segundo tiempo del partido de ida, aumentó considerablemente el número de pases realizados, ya que pasó de hacer una media de 23 pases por cada tiempo de la llave, a efectuar 38 pases, a diferencia de Larry Vásquez, que con mayor protagonismo falló una mayor cantidad de pases; Rafael Robayo erró la misma cantidad en ambos tiempos de la disputa (2).
- **Ómar Vásquez (10):** concentraba mucho el juego, cuando salió en el entretiempo del segundo partido la carga fue dispersada entre Larry Vásquez (5) (media de pases por tiempo en la serie 26, segundo tiempo partido vuelta 51, dif = 25 pases más), Rafael Robayo (8) (media en los tres tiempos de la serie 23, segundo tiempo partido vuelta 38 dif = 15). El cambio en el entretiempo del segundo encuentro no se puede explicar por el número de pases errados, ya que erró 2 pases, pero en comparación con el mismo lapso de tiempo con el partido de ida, ejecutó 13 pases menos, lo que indica que estuvo menos participativo en el partido.

Delanteros

- **Carlos Mosquera (27):** fue el jugador que más pases erró en el partido de ida, con un total de 12 fallos, 9 de ellos en el primer tiempo, en el segundo tiempo del partido de vuelta ejecutó 28 pases errando 5 pases.
- **Mauricio Gómez (11):** los intentos de conectar con sus compañeros de equipo no difieren entre los partidos, sin embargo el número de pases errados fue bastante alto en el primer tiempo del partido de vuelta, ya que de 10 pases intentados 7 fueron errados, la mitad de los intentos de pase (5) fueron ejecutados en el periodo del 30-45, de los cuales 4 fueron errados. Este comportamiento se puede entender por la necesidad del equipo de buscar el gol del empate del encuentro, ya que Corinthians hizo un gol al minuto 28.
- **Edis Ibarguen (18):** estuvo más activo en el compromiso de vuelta, recibiendo más veces el balón y realizando más del doble de pases.

Según las zonas y el mapa de campo determinado, se puede visualizar el comportamiento en las figuras 8 y 9.

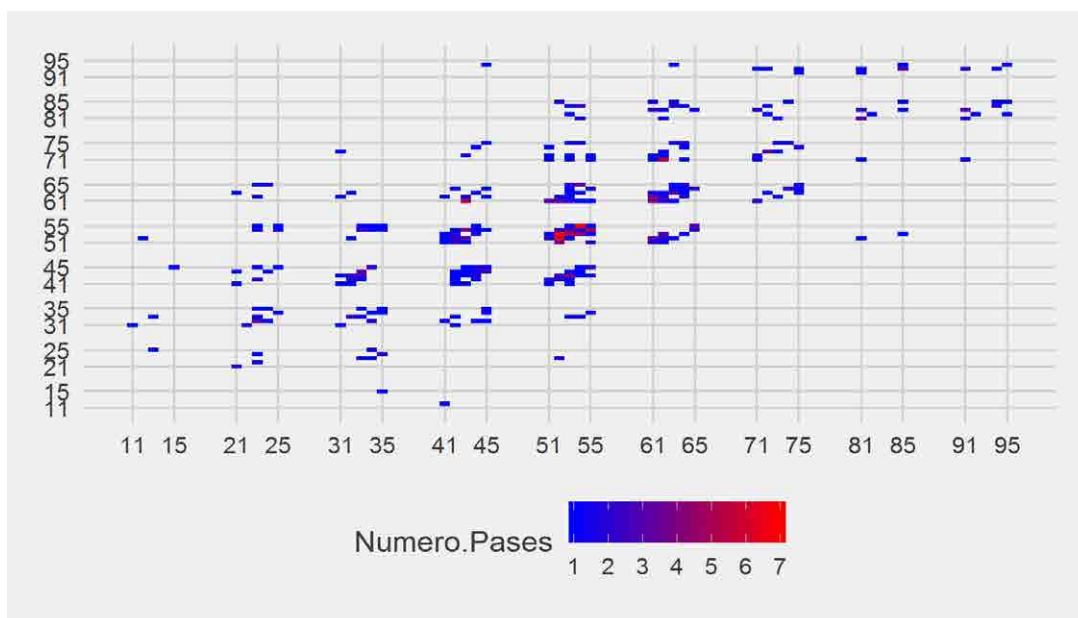


Figura 8. Relación de pases por zona de origen y destino partido de ida.

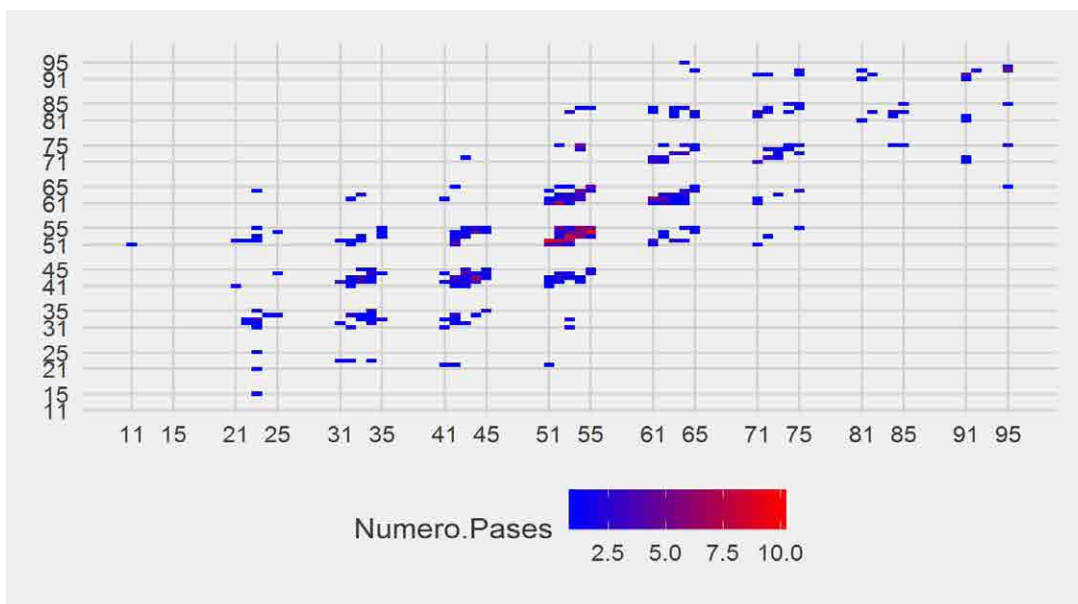


Figura 9. Relación de pases por zona de origen y destino partido de vuelta.

En el partido de ida se realizaron varios pases desde las zonas inferiores de la cancha (12 a 15), se ve una mejor distribución en la circulación del balón entre las zonas 32, 33 y 34 en el partido de vuelta, en el partido de ida se utilizó más el área central de esta zona (33).

Para la zona 4 se ve una tendencia en el partido de ida a realizar más pases en la zona centro-izquierda (zonas: 43, 44, 45) que en la derecha. Este comportamiento no se aprecia en el partido de vuelta, donde el número de pases hacia ese costado se ve disminuido y el juego se recarga netamente en la parte central (42).

En la zona 5 hubo una tendencia a realizar pases cortos dentro de las áreas centralizadas. En esta área se empieza a circular el balón en busca de una mejor salida por las bandas, en el partido de ida se nota una preferencia a enviar el balón a la zona derecha, este comportamiento se equilibra en el partido de vuelta. Este comportamiento tiene su origen en que Carlos Mosquera (27) jugó recargado por esa banda todo el primer compromiso, en el segundo tiempo del partido de vuelta jugó en una zona central recargado al costado izquierdo.

En el partido de vuelta, con la necesidad de empatar la serie, se evidencia un mayor lanzamiento de balones cerca de la portería rival, con un incremento en los balones recibidos en la zona 9 y el lanzamiento de centros desde la zona 95 a las zonas centrales del área rival (83, 92 y 93).

En el partido de ida se jugaron balones largos, por ejemplo, de la zona 31 a la 73 o de la zona 45 a la 94, sin embargo, en el partido de vuelta se evidencia la intención de asegurar la posesión del balón, sin realizar pases con alto riesgo de pérdida.

Análisis partido de vuelta

Arquero

En el partido de ida, en los últimos 15 minutos Álvaro Villete realizó 5 pases, de los cuales 4 fueron errados, estos pases se originaban desde el área propia e iban dirigidos a la zona izquierda del medio campo.

Centrales

En el partido de ida, Danilo Arboleda (3) ejecutó más pases en zonas más adelantadas que su compañero de saga Óscar Cabezas (26). Para el partido de vuelta este comportamiento cambió, ya que Cabezas (26) ejecutó un número considerable de pases desde la zona central del terreno, mientras que Danilo Arboleda (3) los realizó cerca del área propia.

Mediocampistas

- **Larry Vásquez (5):** partido de ida: en los primeros 15 minutos Larry Vásquez (5) intentó abrir el juego desde el centro del campo a la banda izquierda. Desde el minuto 15 al 75 recargó su juego al costado derecho. Después del cambio de Mauricio Gómez (11) vuelve a recostarse al lado izquierdo para jugar con el ingresado Carlos Rodríguez (21). Partido de vuelta: se observa en el primer tiempo que Larry Vásquez (5) sigue recargando sus pases hacia la banda derecha, con el cambio de Ómar Vásquez (10) en el entretiempo los pases son más distribuidos por todo el ancho de la cancha.
- **Rafael Robayo (8):** al ser el volante mixto, su cobertura de área en los pases es más distribuida; sin embargo, se distingue una preferencia a jugar cerca de la zona derecha. Para los primeros 15 minutos del primer tiempo del partido de ida, la presencia en el juego es muy baja, tan solo ejecutó dos pases, uno de ellos errado.
- **Ómar Vásquez (10):** en el primer partido, Ómar Vásquez ejerció un dominio en la zona central del terreno durante el minuto 15 al 30 del primer tiempo, con el paso del tiempo la posesión de la pelota fue disminuyendo y se optó por pases largos. Para el segundo partido se observa que Ómar Vásquez (10) tuvo un mayor desplazamiento sobre el terreno de juego en los primeros 15 minutos del encuentro, en las siguientes etapas del compromiso disminuyó el número de pases realizados como la cobertura que tuvo con los mismos, esto puede explicar la sustitución en el entretiempo.
- **Mauricio Gómez (11):** de los pases errados después del primer gol de Corinthians en el partido de vuelta, 2 fueron balones lanzados desde zonas alejadas del área rival.

Los reportes completos de los datos obtenidos se encuentran publicados en internet en las direcciones suministradas en la Tabla 3, para que el lector pueda tener un mayor dinamismo de los datos obtenidos.

Tabla 3. Reporte de pases por zona por jugador

Documento	URL
Partido de vuelta	https://app.powerbi.com/view?r=eyJrIjoiaOTVjZmY4NjEtNWN-jNS00MTMyLTg3YzUtZGYyNTcyMTY1ZmZhIiwidCI6ImM2Z-TIwZTQzLWVhY2QyNDg5ZS1hMjY2LTQxOGY1NmRmMTI0YyIsImMiOiR9
Partido de ida	https://app.powerbi.com/view?r=eyJrIjoiaNzAyYzQ0ZmEtZWFl-MS00NGY5LWVhY2QyNDg5ZS1hMjY2LTQxOGY1NmRmMTI0YyIsImMiOiR9

Conclusiones

Por medio de las visualizaciones y los reportes generados se puede hacer un análisis de forma fácil y rápida, ya que los datos se encuentran presentados en su contexto, lo cual facilita a las personas que van a tomar decisiones sobre el equipo (cuerpo técnico, entrenadores) la interacción con los datos.

Se ha generado una metodología para el análisis de los pases de jugadores de fútbol a través de herramientas de uso abierto.

Al aplicar la metodología descrita en el documento se pueden beneficiar equipos que no cuentan con la capacidad económica que les permita comprar información y análisis generados por terceros.

A través del uso de un análisis descriptivo se logra adquirir un mayor conocimiento del equipo propio, y a su vez puede servir para entender al rival, con el fin de tomar decisiones en búsqueda de un mejor resultado deportivo.

Al recopilar datos de más partidos, el análisis descriptivo permitirá aplicar técnicas de minería de datos que puedan predecir el comportamiento de un jugador bajo determinadas condiciones. También se recomienda definir un atributo que establezca cómo se terminó perdiendo el balón para precisar secuencias de pases y determinar acciones que faciliten aumentar la posesión del balón.

Para investigaciones futuras se puede recopilar información de los partidos jugados por un equipo en la fase de grupos de un torneo continental, pues esto permite un buen análisis ya que se juegan partidos de ida y vuelta, pero contra tres equipos diferentes, lo que daría más valor al encontrar patrones que generalicen la forma de juego de un equipo dado.

Referencias

- BBC Mundo (2014). ¿Fue el cambio de arquero de Holanda la jugada del Mundial?. *BBC News*.
- Cotta, C., Mora, A. M., Merelo, J. J., & Merelo-Molina, C. (2013). FIFA World Cup 2010: A Network Analysis of the Champion Team Play. *Journal of Systems Science and Complexity*, 26(1), 21-42. DOI: <https://doi.org/10.1007/s11424-013-2291-2>.
- González-Pinzón, M., & González-Sanabria, J. (2013). Aplicación del estándar ISO/IEC 9126-3 en el modelo de datos conceptual entidad-relación. *Revista Facultad de Ingeniería*, 22(35), 113-125. DOI: <https://doi.org/10.19053/01211129.2519>.
- López, J., & Touchette, H. (2013). A Network Theory Analysis of Football Strategies. In *Euromech Physics of Sports Conference*, (pp. 517-528). Palaiseau: Editions de l'Ecole Polytechnique.
- Morenilla, J. (2014). Los penaltis son una ciencia. *El País*.