



# Gol-nectados: La Red del Fútbol Colombiano

Germán Camilo Vásquez Herrera

Facultad de Ciencias  
Estadística  
Universidad Nacional de Colombia  
Bogotá, Colombia  
5 de Marzo de 2025

# Gol-nectados: La Red del Fútbol Colombiano

Germán Camilo Vásquez Herrera

Análisis Estadístico de Redes Sociales - Proyecto

**Profesor:**

Prof. Dr. Juan Camilo Sosa Martínez  
Profesor Asociado - Estadística  
Facultad de Ciencias  
Universidad Nacional de Colombia

Facultad de Ciencias  
Estadística  
Universidad Nacional de Colombia  
Bogotá, Colombia  
5 de Marzo de 2025

# Resumen

## **Gol-nectados: La Red del Fútbol Colombiano**

Este estudio analiza las dinámicas del fútbol colombiano mediante el uso de análisis de redes, modelos de espacio latente y técnicas de clustering. Se construyeron redes para cinco categorías principales (goles, faltas, tarjetas amarillas, tarjetas rojas y goles finales) utilizando datos de las temporadas 2013-1 a 2024-2, con el objetivo de identificar patrones de interacción y evolución temporal entre los equipos. Los resultados revelaron que equipos históricamente relevantes, como **Atlético Nacional** y **Millonarios**, tienen un impacto significativo en la generación de goles y son objetivos frecuentes de sus rivales. Además, se observó una disminución general en la frecuencia de faltas y tarjetas amarillas, lo que sugiere cambios en el estilo de juego. El modelo de espacio latente de dos dimensiones permitió identificar patrones generales de interacción, aunque presentó limitaciones para capturar propiedades estructurales más complejas. Finalmente, se aplicaron técnicas de clustering para identificar comunidades de equipos con comportamientos similares. Este trabajo proporciona insights valiosos para entrenadores, analistas y directivos, destacando la importancia de las conexiones estratégicas y la evolución de las dinámicas competitivas en el fútbol colombiano.

**Palabras clave:** Análisis de redes, modelos de espacio latente, clustering, fútbol colombiano, dinámicas competitivas, interacciones entre equipos.

# Abstract

## Goal-nected: The Colombian Soccer Network

This study analyzes the dynamics of Colombian football using network analysis, latent space models, and clustering techniques. Networks were constructed for five main categories (goals, fouls, yellow cards, red cards, and late goals) using data from the 2013-1 to 2024-2 seasons, with the aim of identifying interaction patterns and temporal evolution among teams. The results revealed that historically relevant teams, such as **Atlético Nacional** and **Millonarios**, have a significant impact on goal generation and are frequent targets of their rivals. Additionally, a general decrease in the frequency of fouls and yellow cards was observed, suggesting changes in playing style. The two-dimensional latent space model allowed for the identification of general interaction patterns, although it had limitations in capturing more complex structural properties. Finally, clustering techniques were applied to identify communities of teams with similar behaviors. This work provides valuable insights for coaches, analysts, and managers, highlighting the importance of strategic connections and the evolution of competitive dynamics in Colombian football.

**Keywords:** Network analysis, latent space models, clustering, Colombian football, competitive dynamics, team interactions.

# Listado de figuras

<b>3-1</b>	Red Goles 2024-2 . . . . .	9
<b>3-2</b>	Red Faltas 2024-2. . . . .	9
<b>3-3</b>	Red Amarillas 2024-2 . . . . .	9
<b>3-4</b>	Red Rojas 2024-2 . . . . .	9
<b>3-5</b>	Red Goles Final 2024-2. . . . .	9
<b>3-6</b>	Evolución de grados de entrada en la red de Goles en el periodo 2013-1 a 2024-2 . . . . .	11
<b>3-7</b>	Evolución de grados de salida en la red de Goles en el periodo 2013-1 a 2024-2. . . . .	11
<b>3-8</b>	Evolución de Goles recibidos en la red de Goles en el periodo 2013-1 a 2024-2 . . . . .	11
<b>3-9</b>	Evolución de Goles realizados en la red de Goles en el periodo 2013-1 a 2024-2. . . . .	11
<b>3-10</b>	Evolución de la centralidad de cercanía en la red de Goles en el periodo 2013-1 a 2024-2. . . . .	11
<b>3-11</b>	Evolución de la centralidad de intermediación en la red de Goles en el periodo 2013-1 a 2024-2	11
<b>3-12</b>	Evolución de la centralidad propia en la red de Goles en el periodo 2013-1 a 2024-2 . . . . .	12
<b>3-13</b>	Evolución de las comunidades en la red de Goles en el periodo 2013-1 a 2024-2 . . . . .	12
<b>3-14</b>	Evolución de la densidad en la red de Goles en el periodo 2013-1 a 2024-2 . . . . .	12
<b>3-15</b>	Evolución de la persistencia en la red de Goles en el periodo 2013-1 a 2024-2 . . . . .	12
<b>3-16</b>	Evolución de grados de entrada en la red de Faltas en el periodo 2013-1 a 2024-2 . . . . .	14
<b>3-17</b>	Evolución de grados de salida en la red de Faltas en el periodo 2013-1 a 2024-2 . . . . .	14
<b>3-18</b>	Evolución de Faltas recibidos en la red de Faltas en el periodo 2013-1 a 2024-2 . . . . .	14
<b>3-19</b>	Evolución de Faltas realizados en la red de Faltas en el periodo 2013-1 a 2024-2. . . . .	14
<b>3-20</b>	Evolución de la centralidad de cercanía en la red de Faltas en el periodo 2013-1 a 2024-2 . . . . .	14
<b>3-21</b>	Evolución de la centralidad de intermediación en la red de Faltas en el periodo 2013-1 a 2024-2	14
<b>3-22</b>	Evolución de la centralidad propia en la red de Faltas en el periodo 2013-1 a 2024-2 . . . . .	15
<b>3-23</b>	Evolución de las comunidades en la red de Faltas en el periodo 2013-1 a 2024-2 . . . . .	15
<b>3-24</b>	Evolución de la densidad en la red de Faltas en el periodo 2013-1 a 2024-2. . . . .	15
<b>3-25</b>	Evolución de la persistencia en la red de Faltas en el periodo 2013-1 a 2024-2 . . . . .	15
<b>3-26</b>	Evolución de grados de entrada en la red de Amarillas en el periodo 2013-1 a 2024-2 . . . . .	16
<b>3-27</b>	Evolución de grados de salida en la red de Amarillas en el periodo 2013-1 a 2024-2. . . . .	16
<b>3-28</b>	Evolución de Amarillas recibidos en la red de Amarillas en el periodo 2013-1 a 2024-2 . . . . .	16
<b>3-29</b>	Evolución de Amarillas realizados en la red de Amarillas en el periodo 2013-1 a 2024-2 . . . . .	16
<b>3-30</b>	Evolución de la centralidad de cercanía en la red de Amarillas en el periodo 2013-1 a 2024-2 . . . . .	16
<b>3-31</b>	Evolución de la centralidad de intermediación en la red de Amarillas en el periodo 2013-1 a 2024-2	16
<b>3-32</b>	Evolución de la centralidad propia en la red de Amarillas en el periodo 2013-1 a 2024-2 . . . . .	17
<b>3-33</b>	Evolución de las comunidades en la red de Amarillas en el periodo 2013-1 a 2024-2 . . . . .	17

## Gol-nectados: La Red del Fútbol Colombiano

---

<b>3-34</b>	Evolución de la densidad en la red de Amarillas en el periodo 2013-1 a 2024-2 . . . . .	17
<b>3-35</b>	Evolución de la persistencia en la red de Amarillas en el periodo 2013-1 a 2024-2 . . . . .	17
<b>3-36</b>	Evolución de grados de entrada en la red de Rojas en el periodo 2013-1 a 2024-2 . . . . .	18
<b>3-37</b>	Evolución de grados de salida en la red de Rojas en el periodo 2013-1 a 2024-2 . . . . .	18
<b>3-38</b>	Evolución de Rojas recibidos en la red de Rojas en el periodo 2013-1 a 2024-2 . . . . .	18
<b>3-39</b>	Evolución de Rojas realizados en la red de Rojas en el periodo 2013-1 a 2024-2 . . . . .	18
<b>3-40</b>	Evolución de la centralidad de cercanía en la red de Rojas en el periodo 2013-1 a 2024-2. . .	18
<b>3-41</b>	Evolución de la centralidad de intermediación en la red de Rojas en el periodo 2013-1 a 2024-2	18
<b>3-42</b>	Evolución de la centralidad propia en la red de Rojas en el periodo 2013-1 a 2024-2 . . . . .	19
<b>3-43</b>	Evolución de las comunidades en la red de Rojas en el periodo 2013-1 a 2024-2 . . . . .	19
<b>3-44</b>	Evolución de la densidad en la red de Rojas en el periodo 2013-1 a 2024-2 . . . . .	19
<b>3-45</b>	Evolución de la persistencia en la red de Rojas en el periodo 2013-1 a 2024-2 . . . . .	19
<b>3-46</b>	Evolución de grados de entrada en la red de Goles Final en el periodo 2013-1 a 2024-2 . . .	20
<b>3-47</b>	Evolución de grados de salida en la red de Goles Final en el periodo 2013-1 a 2024-2 . . . .	20
<b>3-48</b>	Evolución de Goles Final recibidos en la red de Goles Final en el periodo 2013-1 a 2024-2 . .	20
<b>3-49</b>	Evolución de Goles Final realizados en la red de Goles Final en el periodo 2013-1 a 2024-2 . .	20
<b>3-50</b>	Evolución de la centralidad de cercanía en la red de Goles Final en el periodo 2013-1 a 2024-2	20
<b>3-51</b>	Evolución de la centralidad de intermediación en la red de Goles Final en el periodo 2013-1 a 2024-2 . . . . .	20
<b>3-52</b>	Evolución de la centralidad propia en la red de Goles Final en el periodo 2013-1 a 2024-2 . .	21
<b>3-53</b>	Evolución de las comunidades en la red de Goles Final en el periodo 2013-1 a 2024-2 . . . . .	21
<b>3-54</b>	Evolución de la densidad en la red de Goles Final en el periodo 2013-1 a 2024-2. . . . .	21
<b>3-55</b>	Evolución de la persistencia en la red de Goles Final en el periodo 2013-1 a 2024-2 . . . . .	21
<b>3-56</b>	Convergencia Goles 2024-2 . . . . .	22
<b>3-57</b>	Convergencia Faltas 2024-2. . . . .	22
<b>3-58</b>	Convergencia Amarillas 2024-2 . . . . .	22
<b>3-59</b>	Convergencia Rojas 2024-2 . . . . .	22
<b>3-60</b>	Convergencia Goles Final 2024-2 . . . . .	22
<b>3-61</b>	Inferencia sobre el intercepto Goles 2024-2 . . . . .	23
<b>3-62</b>	Inferencia sobre el intercepto Faltas 2024-2 . . . . .	23
<b>3-63</b>	Inferencia sobre el intercepto Amarillas 2024-2. . . . .	23
<b>3-64</b>	Inferencia sobre el intercepto Rojas 2024-2 . . . . .	23
<b>3-65</b>	Inferencia sobre el intercepto Goles Final 2024-2 . . . . .	23
<b>3-66</b>	Inferencia sobre las posiciones latentes Goles 2024-2 . . . . .	25
<b>3-67</b>	Inferencia sobre las posiciones latentes Faltas 2024-2 . . . . .	25
<b>3-68</b>	Inferencia sobre las posiciones latentes Amarillas 2024-2 . . . . .	25
<b>3-69</b>	Inferencia sobre las posiciones latentes Rojas 2024-2 . . . . .	25
<b>3-70</b>	Inferencia sobre las posiciones latentes Goles Final 2024-2. . . . .	25
<b>3-71</b>	Inferencia probabilidades de interacción Goles 2024-2 . . . . .	26
<b>3-72</b>	Inferencia probabilidades de interacción Faltas 2024-2 . . . . .	26
<b>3-73</b>	Inferencia probabilidades de interacción Amarillas 2024-2 . . . . .	26
<b>3-74</b>	Inferencia probabilidades de interacción Rojas 2024-2 . . . . .	26
<b>3-75</b>	Inferencia probabilidades de interacción Goles Final 2024-2 . . . . .	26
<b>3-76</b>	Bondad de ajuste Goles 2024-2 . . . . .	27

---

**Gol-nectados: La Red del Fútbol Colombiano**

---

<b>3-77</b>	Bondad de ajuste Faltas 2024-2 . . . . .	27
<b>3-78</b>	Bondad de ajuste Amarillas 2024-2 . . . . .	27
<b>3-79</b>	Bondad de ajuste Rojas 2024-2 . . . . .	27
<b>3-80</b>	Bondad de ajuste Goles Final 2024-2 . . . . .	27
<b>3-81</b>	Clustering de posiciones latentes Goles 2024-2 . . . . .	29
<b>3-82</b>	Clustering de posiciones latentes Faltas 2024-2 . . . . .	29
<b>3-83</b>	Clustering de posiciones latentes Amarillas 2024-2 . . . . .	29
<b>3-84</b>	Clustering de posiciones latentes Rojas 2024-2 . . . . .	29
<b>3-85</b>	Clustering de posiciones latentes Goles Final 2024-2 . . . . .	29

# Lista de tablas

# Contenido

<b>Resumen</b>	<b>II</b>
<b>Abstract</b>	<b>III</b>
<b>Lista de figuras</b>	<b>IV</b>
<b>Lista de tablas</b>	<b>VII</b>
<b>Contenido</b>	<b>VIII</b>
<b>1 Introducción</b>	<b>1</b>
<b>2 Metodología</b>	<b>3</b>
2.1 Marco Teórico . . . . .	3
2.1.1 Análisis de Redes . . . . .	3
2.1.2 Modelos de Espacio Latente . . . . .	3
2.1.3 Clustering en Redes . . . . .	5
2.1.4 Aplicaciones de Redes en Deportes . . . . .	5
2.1.5 Integración de Métodos . . . . .	5
2.2 Método . . . . .	6
<b>3 Aplicación</b>	<b>8</b>
3.1 Visualización de redes . . . . .	8
3.2 Análisis Longitudinal . . . . .	10
3.2.1 Goles . . . . .	11
3.2.2 Faltas . . . . .	13
3.2.3 Amarillas . . . . .	16
3.2.4 Rojas . . . . .	18
3.2.5 Goles Final . . . . .	20
3.3 Modelo de espacio latente de dos dimensiones . . . . .	22
3.3.1 Convergencia . . . . .	22
3.3.2 Inferencia sobre el intercepto . . . . .	23
3.3.3 Inferencia sobre las posiciones latentes . . . . .	24
3.3.4 Inferencia probabilidades de interacción . . . . .	26
3.3.5 Bondad de ajuste . . . . .	27
3.3.6 Clustering de posiciones latentes . . . . .	28

---

**Gol-nectados: La Red del Fútbol Colombiano**

3.3.7 Proporción de los equipos en las comunidades (Cluster) . . . . .	30
<b>4 Discusión de resultados</b>	<b>35</b>
<b>5 Conclusiones</b>	<b>37</b>
<b>6 Recomendaciones</b>	<b>38</b>
<b>Referencias Bibliográficas</b>	<b>39</b>

# 1 Introducción

El análisis de redes se ha consolidado como una herramienta fundamental para estudiar sistemas complejos en diversos campos, desde las ciencias sociales hasta la biología y el deporte. En particular, su aplicación en el ámbito deportivo ha permitido analizar dinámicas de interacción entre equipos y jugadores, identificando patrones ocultos y facilitando la toma de decisiones estratégicas. En el contexto del fútbol, el análisis de redes ha sido utilizado para examinar desde la eficacia de los pases hasta la identificación de jugadores clave, ofreciendo una perspectiva integral sobre el rendimiento y la estrategia de los equipos. Este estudio se enfoca en el fútbol colombiano, utilizando técnicas de análisis de redes para explorar las dinámicas de interacción entre equipos a lo largo de múltiples temporadas, con el objetivo de identificar patrones de comportamiento y su evolución en el tiempo.

El marco teórico de este trabajo se sustenta en tres pilares principales: el análisis de redes, los modelos de espacio latente y el clustering. El análisis de redes permite representar y analizar las interacciones entre equipos a través de métricas como el grado, la centralidad y la densidad, las cuales han demostrado ser útiles para entender la estructura y dinámica de las relaciones en el deporte. Por su parte, los modelos de espacio latente ofrecen una representación de baja dimensionalidad de las interacciones, asumiendo que la probabilidad de conexión entre nodos depende de su proximidad en un espacio social latente. Finalmente, el clustering permite identificar comunidades dentro de las redes, lo que facilita la detección de grupos con comportamientos similares. La integración de estas metodologías proporciona una visión holística de las dinámicas del fútbol colombiano, permitiendo no solo describir la estructura de las interacciones, sino también predecir comportamientos futuros .

El objetivo principal de este estudio es analizar las dinámicas del fútbol colombiano a través de redes construidas para cinco categorías principales: goles, faltas, tarjetas amarillas, tarjetas rojas y goles finales. Estas redes se construyen a partir de datos de las temporadas 2013-1 a 2024-2, centrándose en los equipos con mayor relevancia histórica en el fútbol colombiano. Además, se busca identificar patrones ocultos en las interacciones entre equipos mediante el ajuste de un modelo de espacio latente y la aplicación de técnicas de clustering. Este enfoque permite no solo describir la estructura de las redes, sino también explorar su evolución a lo largo del tiempo, ofreciendo insights valiosos para entrenadores, analistas y aficionados.

La justificación de este trabajo radica en la necesidad de aplicar metodologías cuantitativas avanzadas para entender las dinámicas del fútbol colombiano, un deporte que tiene un impacto cultural y social significativo en el país. Aunque existen estudios previos que han aplicado el análisis de redes al fútbol, la mayoría se han centrado en ligas europeas o en aspectos específicos del juego, como los pases o los tiros. Este estudio contribuye a la literatura al ofrecer un análisis integral de múltiples categorías y al aplicar modelos de espacio latente y clustering para identificar patrones ocultos en las interacciones entre equipos. Además, los resultados de este trabajo tienen el potencial de informar decisiones estratégicas y tácticas en el fútbol colombiano, mejorando así el rendimiento de los equipos y enriqueciendo la experiencia de los aficionados.

El alcance del proyecto abarca el análisis de 115 redes construidas a partir de datos de 23 temporadas, lo que permite una visión longitudinal de las dinámicas del fútbol colombiano. Se utilizan técnicas de análisis

---

## 1. Introducción

de redes, modelos de espacio latente y clustering para explorar las interacciones entre equipos, identificando patrones y tendencias a lo largo del tiempo. Además, se generan visualizaciones interactivas y tablas que facilitan la interpretación de los resultados, los cuales están disponibles en un archivo [HTML](#) de libre acceso. Este enfoque no solo proporciona una comprensión profunda de las dinámicas del fútbol colombiano, sino que también establece un marco metodológico que puede ser aplicado a otros deportes y contextos.

## 2 Metodología

### 2.1 Marco Teórico

El análisis de redes se ha convertido en una herramienta fundamental para estudiar interacciones complejas en diversos campos, desde las ciencias sociales hasta la biología y el deporte. En este estudio, se utiliza el análisis de redes para examinar las dinámicas del fútbol colombiano, aplicando conceptos teóricos y metodológicos que han sido ampliamente desarrollados en la literatura. A continuación, se presenta una revisión de los conceptos clave que sustentan este trabajo.

#### 2.1.1 Análisis de Redes

El análisis de redes es una metodología que permite representar y analizar relaciones entre entidades (nodos) a través de conexiones (aristas). Esta técnica ha sido utilizada para estudiar sistemas complejos en diversas disciplinas *Newman* [2010]. En el contexto del deporte, las redes permiten modelar interacciones entre equipos o jugadores, identificando patrones y estructuras ocultas *Cotta et al.* [2013].

Las métricas de redes, como el grado, la centralidad y la densidad, son fundamentales para entender la estructura y dinámica de las interacciones. Por ejemplo, el grado de un nodo indica cuántas conexiones tiene, mientras que la centralidad mide su influencia dentro de la red *Freeman* [1979]. Estas métricas han sido aplicadas en estudios deportivos para analizar el rendimiento de equipos y jugadores *Grund* [2012].

#### 2.1.2 Modelos de Espacio Latente

Los modelos de espacio latente son una técnica estadística que permite representar relaciones entre nodos en un espacio de baja dimensionalidad. Estos modelos asumen que la probabilidad de una conexión entre dos nodos depende de su proximidad en este espacio latente *Hoff et al.* [2002a]. Esta metodología ha sido utilizada para analizar redes sociales, biológicas y deportivas, identificando patrones ocultos en las interacciones *Handcock et al.* [2007]. En el contexto del deporte, los modelos de espacio latente han sido aplicados para estudiar la dinámica de equipos y jugadores, permitiendo identificar grupos con comportamientos similares *Gormley & Murphy* [2010]. Además, estos modelos son útiles para predecir interacciones futuras basadas en patrones históricos *Kravitsky & Handcock* [2009].

Una contribución clave en el modelado de datos relacionales es la inclusión de variables latentes, las cuales representan atributos no observables de los nodos. Aunque estas variables no son directamente medibles, son fundamentales para determinar las probabilidades de interacción en el grafo. En este marco, las entradas

de la matriz de adyacencia  $Y = [y_{i,j}]$ , que describe el grafo  $G = (V, E)$ , se modelan como independientes condicionalmente, siguiendo una distribución Bernoulli de la forma:

$$y_{i,j} \mid \mu, \beta, x_{i,j}, u_i, u_j \stackrel{\text{ind}}{\sim} \text{Bernoulli}\left(g\left(\mu + \beta^T x_{i,j} + \alpha(u_i, u_j)\right)\right),$$

donde:

- $g(\cdot)$  es una función de enlace, que puede ser la función logit inversa,  $\text{expit}(x)$ , o la función de distribución acumulativa de una Normal estándar,  $\Phi(x)$ .
- $\mu$  es el término de intercepto.
- $\beta = (\beta_1, \dots, \beta_p)$  es un vector de coeficientes asociados a las covariables (variables exógenas)  $x_{i,j}$ .
- $x_{i,j} = (x_{i,j,1}, \dots, x_{i,j,p})$  es un vector de covariables derivadas de las características de los nodos.
- $u_i = (u_{i,1}, \dots, u_{i,k})$  es un vector de variables latentes ubicadas en un espacio social de dimensión  $k$ .
- $\alpha(\cdot, \cdot)$  es una función simétrica que modela las interacciones latentes entre nodos, con posibles formulaciones:
  - Modelo de distancia:  $\alpha(u_i, u_j) = -\|u_i - u_j\|$ .
  - Modelo bilineal:  $\alpha(u_i, u_j) = u_i^T u_j$ .
  - Modelo factorial:  $\alpha(u_i, u_j) = u_i^T \Lambda u_j$ , donde  $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_k)$  es una matriz diagonal de tamaño  $k \times k$ .

Para la implementación de este modelo, es necesario definir:

- La dimensión del espacio latente  $k$ .
- La distribución de las variables latentes.
- La forma funcional de  $\alpha(\cdot, \cdot)$ .
- La función de enlace  $g(\cdot)$ .

Debido a la naturaleza jerárquica de estos modelos, un enfoque Bayesiano es especialmente adecuado. Este enfoque permite incorporar incertidumbre y asignar distribuciones a priori a los parámetros, utilizando métodos de inferencia como las cadenas de Markov de Monte Carlo (MCMC) o técnicas variacionales, que son eficaces para explorar la distribución posterior de los parámetros y las variables latentes.

## Modelo de distancia

El modelo de distancia básico se expresa como:

$$y_{i,j} \mid \mu, u_i, u_j \stackrel{\text{ind}}{\sim} \text{Bernoulli}\left(g\left(\mu - \|u_i - u_j\|\right)\right),$$

donde:

- $g(x)$  es la función de enlace, comúnmente la función logit inversa ( $\text{expit}(x)$ ):

$$\text{expit}(x) = \text{logit}^{-1}(x) = \frac{1}{1 + \exp(-x)} = \frac{\exp(x)}{1 + \exp(x)}.$$

- $\mu$  es el intercepto global, con una distribución a priori:

$$\mu | \xi, \psi^2 \sim \mathcal{N}(\xi, \psi^2),$$

donde  $\xi$  y  $\psi^2$  son hiperparámetros que regulan la media y la varianza del intercepto.

- $u_i$  representa las posiciones latentes de los nodos en un espacio euclíadiano de dimensión  $k$ , con una distribución a priori:

$$u_i | \mu, \sigma^2 \stackrel{\text{ind}}{\sim} \mathcal{N}_k(\mu, \sigma^2 I),$$

donde  $\mu$  es el centro del espacio latente y  $\sigma^2$  controla la dispersión de las posiciones.

Este modelo, propuesto inicialmente por *Hoff et al.* [2002b], es ampliamente empleado en el análisis de redes sociales para modelar la proximidad latente entre nodos, donde la probabilidad de conexión disminuye conforme aumenta la distancia euclíadiana entre sus posiciones latentes. En el ámbito deportivo, su aplicación permite identificar patrones de interacción entre jugadores y equipos, así como predecir comportamientos futuros basados en su cercanía en el espacio latente.

### 2.1.3 Clustering en Redes

El clustering es una técnica que permite identificar comunidades o grupos dentro de una red, donde los nodos dentro de un mismo grupo están más conectados entre sí que con nodos de otros grupos *Fortunato* [2010]. Algoritmos como *walktrap* y *k-means* son ampliamente utilizados para este propósito *Pons & Latapy* [2005].

En el deporte, el clustering ha sido aplicado para analizar la estructura de equipos y la dinámica de juegos. Por ejemplo, se ha utilizado para identificar estrategias de juego y roles específicos de los jugadores *Clemente et al.* [2015]. Además, el clustering permite estudiar la evolución de las interacciones a lo largo del tiempo, identificando cambios en la estructura de la red *Lucey et al.* [2013].

### 2.1.4 Aplicaciones de Redes en Deportes

El análisis de redes ha sido ampliamente utilizado en el estudio de deportes, desde el fútbol hasta el baloncesto y el tenis. Estas aplicaciones incluyen el análisis de pases entre jugadores, la identificación de jugadores clave y la evaluación de estrategias de equipo *Gyarmati et al.* [2014].

En el fútbol, por ejemplo, las redes han sido utilizadas para estudiar la eficacia de los pases y la coordinación entre jugadores *Clemente et al.* [2016]. En otros deportes, como el baloncesto, las redes han permitido analizar la dinámica de los tiros y la defensa *Livingston* [2012]. Estas aplicaciones demuestran la versatilidad del análisis de redes para entender las dinámicas deportivas.

### 2.1.5 Integración de Métodos

La combinación de análisis de redes, modelos de espacio latente y clustering ofrece una perspectiva integral para estudiar sistemas complejos. Esta integración permite no solo describir la estructura de las interacciones, sino también identificar patrones ocultos y predecir comportamientos futuros *Hoff* [2008]. En el contexto del deporte, esta metodología ha sido utilizada para analizar la evolución de equipos y la efectividad de estrategias *Petersen-Wagner* [2019].

En resumen, el marco teórico presentado sustenta el uso del análisis de redes, modelos de espacio latente y clustering para estudiar las dinámicas del fútbol colombiano. Estas metodologías, respaldadas por una amplia literatura, permiten una comprensión profunda de las interacciones entre equipos y su evolución a lo largo del tiempo.

## 2.2 Método

Este estudio se enmarca en un enfoque cuantitativo basado en el análisis de redes, con el objetivo de examinar las dinámicas del fútbol colombiano a través de cinco categorías principales: goles, faltas, tarjetas amarillas, tarjetas rojas y goles finales (goles marcados entre el minuto 76 y el final del partido). Se utilizaron datos de las temporadas 2013-1 a 2024-2, centrándose en los equipos con mayor relevancia histórica en el fútbol colombiano.

Los datos fueron recolectados de fuentes oficiales, incluyendo registros públicos de la liga colombiana y bases de datos especializadas en estadísticas deportivas. Se seleccionaron únicamente los partidos de la primera división, excluyendo torneos internacionales y categorías inferiores. Los datos incluyen información detallada sobre goles, faltas, tarjetas y otros eventos relevantes durante los partidos.

Una vez recolectados, los datos fueron sometidos a un proceso de limpieza para eliminar registros duplicados y corregir errores. Posteriormente, se transformaron en matrices de adyacencia para cada categoría, donde los nodos representan equipos y las aristas representan interacciones entre ellos (por ejemplo, goles anotados, faltas cometidas, etc.). Estas matrices tienen en cuenta el peso (por ejemplo, la cantidad de goles hechos y recibidos) y la direccionalidad de las conexiones, lo que permite construir redes ponderadas y dirigidas.

En total, se construyeron 115 redes, divididas en 23 temporadas para las cinco categorías mencionadas. En cada red, el tamaño de los nodos se determinó en función de las conexiones de salida y su peso (por ejemplo, a cuántos equipos se les marcó y cuántos goles se anotaron). Además, se asignó un color específico a los nodos que representan equipos con al menos un título de primera división en Colombia, lo que facilita la identificación visual de los equipos más exitosos.

Para estudiar la dinámica y evolución del comportamiento de los equipos y las redes a través del tiempo, se aplicó un análisis longitudinal a diversas métricas, incluyendo técnicas de centralidad (grado, cercanía, intermediación), detección de comunidades utilizando el algoritmo *cluster\_walktrap*, densidad y persistencia de conexiones. Este análisis permitió identificar patrones y tendencias en las interacciones entre equipos a lo largo de las temporadas.

Además, se ajustó un modelo de espacio latente de dos dimensiones a cada una de las 115 redes para identificar patrones ocultos en las interacciones entre equipos. Cabe destacar que, debido a la alta densidad de las redes de faltas y tarjetas amarillas, el modelo no convergía inicialmente. Para solucionar este problema, se seleccionó un umbral del 50% para las aristas con mayor peso, conservando solo las conexiones más significativas. Este enfoque permitió entrenar y ajustar el modelo de manera efectiva.

El modelo fue validado mediante pruebas de convergencia y bondad de ajuste, utilizando métricas como la densidad, assortatividad y transitividad. También se evaluó la inferencia sobre el intercepto, las posiciones latentes y las probabilidades de interacción para verificar la precisión del modelo en la estimación de estos parámetros y asociaciones.

Para identificar comunidades en función de las posiciones latentes estimadas por el modelo, se realizó un proceso de clustering utilizando el algoritmo *k-means*. Previamente, se ajustó un modelo de mezcla gaussiana para determinar el número óptimo de grupos ( $k$ ) mediante el Criterio de Información Bayesiano (BIC). Este enfoque permitió identificar agrupaciones naturales en los datos y analizar la influencia de cada equipo

## 2. Metodología

### 2.2. Método

dentro de las comunidades detectadas. Se generaron tablas que muestran las proporciones de cada equipo en cada comunidad, lo que facilita la interpretación de los resultados.

Con el fin de facilitar la revisión y el acceso a los resultados, se creó un archivo [HTML](#) que contiene las 855 imágenes y 115 tablas generadas en el estudio. Este archivo es de libre acceso y permite explorar de manera interactiva los resultados detallados. Todas las implementaciones se realizaron en R, utilizando librerías especializadas como `igraph` y `networkx`.

Dado que los datos utilizados son de carácter público y no incluyen información personal sensible, no se requirieron aprobaciones éticas adicionales. Sin embargo, se aseguró el manejo responsable de los datos, respetando las políticas de uso de las fuentes consultadas.

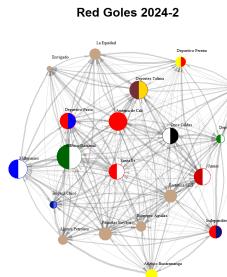
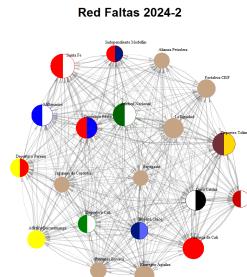
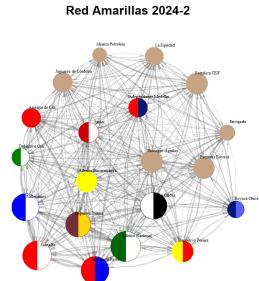
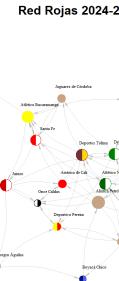
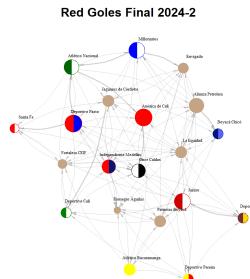
## 3 Aplicación

### 3.1 Visualización de redes

Como se mencionó anteriormente, se construyeron un total de **115 redes**, cada una correspondiente a diferentes aspectos del análisis realizado. En este documento, se presentan y visualizan **5 redes representativas**, las cuales corresponden a la última temporada disponible a la fecha del fútbol colombiano (2024-2) para cada una de las categorías analizadas: **Goles, Faltas, Tarjetas Amarillas, Tarjetas Rojas y Goles Final**. Estas redes fueron seleccionadas para ofrecer una visión clara y concisa de los resultados más relevantes.

Para aquellos interesados en explorar el conjunto completo de redes, así como los resultados y salidas detalladas generadas en este estudio, se ha habilitado un **archivo HTML interactivo**. Puede acceder a él haciendo clic [aquí](#), donde encontrará la visualización de todas las redes construidas, junto con los análisis complementarios y datos adicionales.

A continuación, se presenta la visualización de las redes correspondientes a la **temporada 2024-2** para cada una de las categorías analizadas:

**Figura 3-1:** Red Goles 2024-2**Figura 3-2:** Red Faltas 2024-2**Figura 3-3:** Red Amarillas 2024-2**Figura 3-4:** Red Rojas 2024-2**Figura 3-5:** Red Goles Final 2024-2

Recordemos que el tamaño de los nodos en cada red está determinado por la cantidad de eventos registrados, ya sean goles, faltas, tarjetas amarillas, entre otros. Al observar las redes correspondientes a la temporada 2024-II, se pueden identificar algunos patrones preliminares que merecen atención.

En la categoría de **goles**, los nodos más grandes corresponden a **Atlético Nacional**, **Millonarios** y **Deportes Tolima**, lo que sugiere que estos equipos podrían estar asociados con una mayor cantidad de goles. Por otro lado, los nodos más pequeños corresponden a **Jaguares** y **Envigado**, lo que sugiere que estos equipos marcaron menos goles durante la temporada. Sin embargo, se observa que ambos tienen una mayor cantidad de **conexiones de entrada**, lo que podría indicar que fueron unos de los equipos contra los que más equipos distintos marcaron al menos un gol.

En el caso de las **faltas**, dado que es un aspecto inherente al fútbol, se observa que todos los equipos están conectados entre sí. Sin embargo, los nodos más grandes corresponden a **Santa Fe**, **Atlético Nacional** y **Deportes Tolima**, lo que podría sugerir una mayor cantidad de faltas cometidas por estos equipos.

### 3.2. Análisis Longitudinal

### 3. Aplicación

Respecto a las **tarjetas amarillas**, se observa un comportamiento similar al de las faltas, con una distribución generalizada entre todos los equipos. No obstante, los nodos más grandes corresponden a **Atlético Nacional**, **Santa Fe**, **Millonarios** y **Once Caldas**, lo que podría indicar una mayor acumulación de amonestaciones en estos equipos.

En la categoría de **tarjetas rojas**, aunque los nodos son generalmente más pequeños debido a la menor frecuencia de expulsiones, se destaca que el nodo correspondiente a **Independiente Medellín** no muestra tarjetas rojas recibidas, pero sí una expulsión provocada en su enfrentamiento contra **Atlético Bucaramanga**.

Finalmente, en la red de **goles final**, los nodos más grandes corresponden a **Millonarios**, **América de Cali** y **Deportivo Pasto**, lo que podría sugerir que estos equipos marcaron más goles en los últimos 15 minutos. Además, se observa una conexión destacada entre **Millonarios** y **Atlético Nacional**, quienes se enfrentaron en tres ocasiones durante la temporada, con **Millonarios** anotando 2 de los 4 goles que le marcó a **Atlético Nacional** en los minutos finales.

## 3.2 Análisis Longitudinal

A continuación, se analiza la evolución del comportamiento de las redes a lo largo del periodo **2013-1** a **2024-2**, centrándose en diversas métricas de interés. Para facilitar la visualización y el análisis, se han seleccionado los **nueve equipos con más títulos de liga en Colombia**, ya que suelen ser los de mayor relevancia en este tipo de estudios.

**Tabla 3-1:** Equipos con más títulos de liga en Colombia

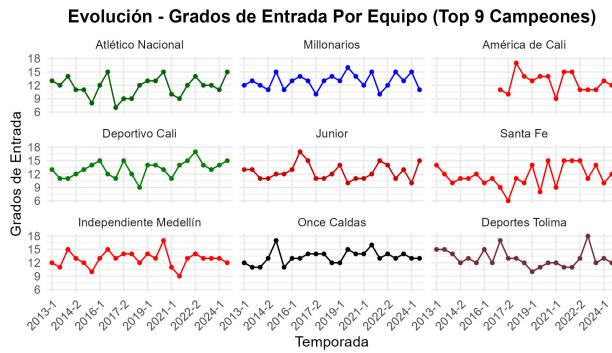
Equipo	Títulos de Liga
Atlético Nacional	18
Millonarios	16
América de Cali	15
Deportivo Cali	10
Junior	10
Santa Fe	9
Independiente Medellín	6
Once Caldas	4
Deportes Tolima	3
Deportes Quindío	1
Unión Magdalena	1
Deportivo Pasto	1
Cúcuta Deportivo	1
Boyacá Chicó	1
Deportivo Pereira	1
Atlético Bucaramanga	1

Cabe destacar que **América de Cali** no participó en la primera división entre 2012 y 2016, al competir en la segunda división del fútbol colombiano. Por lo tanto, como se podrá observar en los gráficos, la información disponible para este equipo comienza a partir de la temporada **2017-1**.

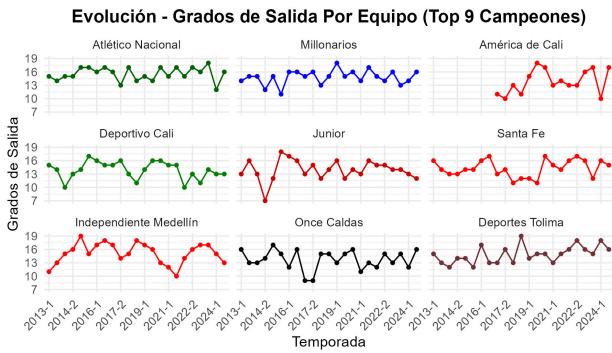
### 3. Aplicación

### 3.2. Análisis Longitudinal

#### 3.2.1 Goles



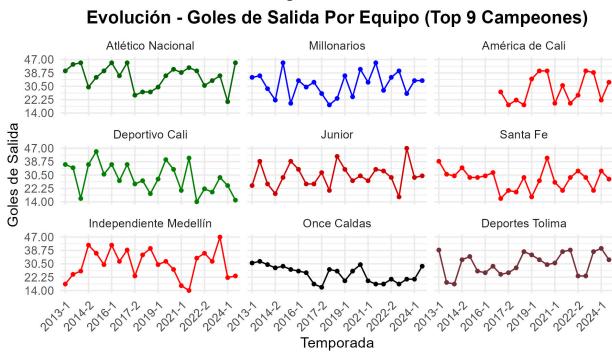
**Figura 3-6:** Evolución de grados de entrada en la red de Goles en el periodo 2013-1 a 2024-2



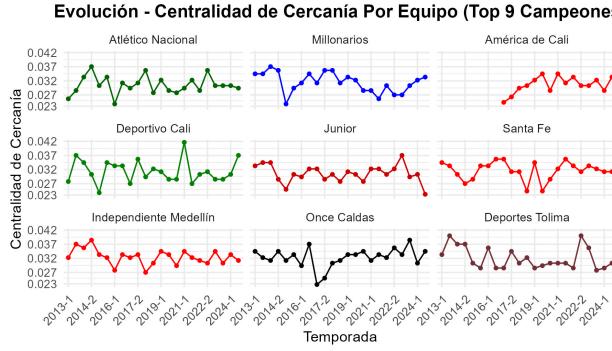
**Figura 3-7:** Evolución de grados de salida en la red de Goles en el periodo 2013-1 a 2024-2



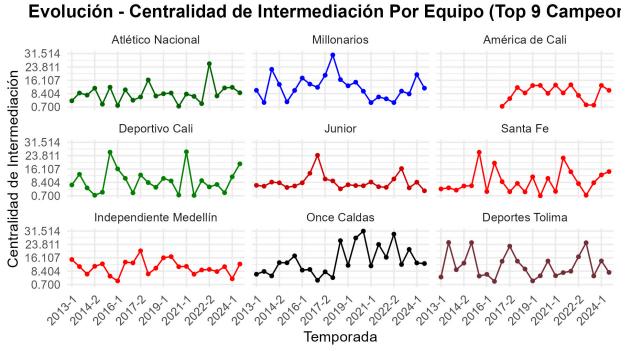
**Figura 3-8:** Evolución de Goles recibidos en la red de Goles en el periodo 2013-1 a 2024-2



**Figura 3-9:** Evolución de Goles realizados en la red de Goles en el periodo 2013-1 a 2024-2



**Figura 3-10:** Evolución de la centralidad de cercanía en la red de Goles en el periodo 2013-1 a 2024-2



**Figura 3-11:** Evolución de la centralidad de intermediación en la red de Goles en el periodo 2013-1 a 2024-2

Al observar la evolución de los grados de entrada (cantidad de equipos que lograron anotarle goles a un equipo), se destaca que la peor defensa fue la de Deportes Tolima en la temporada 2023-1, donde todos los equipos rivales le marcaron al menos un gol. Le siguen América de Cali en 2018-1, Deportivo Cali en 2022-2, Junior en 2016-2, Independiente Medellín en 2020-1, Once Caldas en 2015-1 y nuevamente Deportes Tolima en 2017-1, donde todos sus rivales, excepto uno, lograron anotarles al menos un gol. Por otro lado, la mejor defensa fue la de Santa Fe en 2017-2, ya que solo seis equipos lograron anotarle, incluyendo Millonarios, contra quien perdió la final de la liga. La segunda mejor defensa fue la de Atlético Nacional en 2017-1, donde solo siete equipos le marcaron goles. Se observa una variabilidad en el rendimiento de los equipos, siendo Once Caldas el más estable, ya que recibió goles de entre 11 y 14 equipos en promedio por temporada.

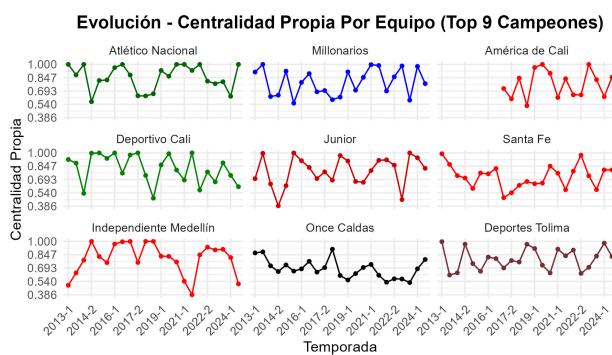
### 3.2. Análisis Longitudinal

### 3. Aplicación

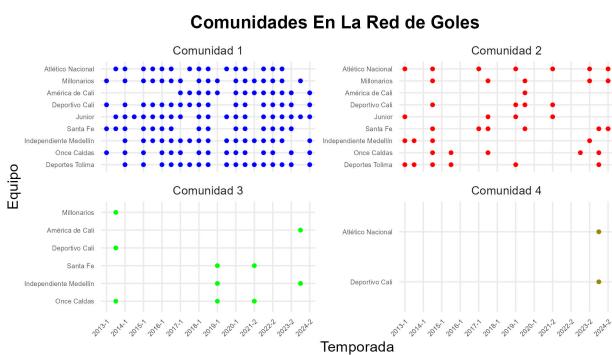
En cuanto a los grados de salida (cantidad de equipos a los que un equipo logró anotar goles), Atlético Nacional es el más regular, manteniendo una estabilidad y anotando a casi todos sus rivales temporada tras temporada. Millonarios sigue en segundo lugar, con un rendimiento ligeramente inferior. Destacan Independiente Medellín en 2015-1 y Deportes Tolima en 2018-2, únicos equipos en el periodo estudiado que lograron anotarle a todos sus rivales. Por el contrario, Junior en 2014-2 fue el peor, anotando solo a siete de los 17 rivales posibles.

En términos de goles recibidos, Millonarios y Junior muestran una estabilidad notable, siendo de los equipos que menos goles recibieron temporada tras temporada. Sin embargo, otros equipos como Deportivo Cali presentan una alta variabilidad, destacando las temporadas 2014-2 y 2023-2, donde recibieron alrededor de 40 goles. En cuanto a los goles convertidos, todos los equipos, excepto Once Caldas, muestran una alta variabilidad. Independiente Medellín y Junior en 2023-2 marcaron 47 goles, el máximo en el periodo analizado. Atlético Nacional, Millonarios y Deportivo Cali también se acercaron a esta cifra en varias temporadas. Once Caldas, por su parte, se caracteriza por su estabilidad pero con pocos goles anotados, rara vez superando los 30 goles por temporada.

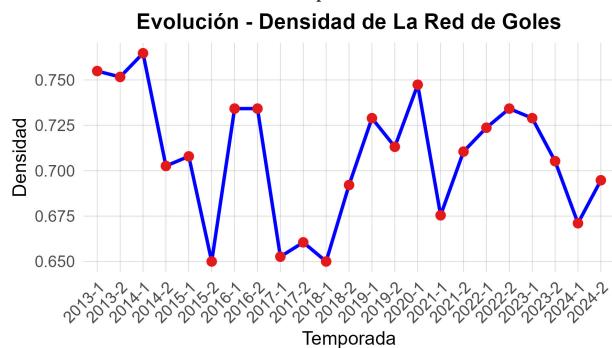
Según la centralidad de cercanía, los equipos más influyentes en la generación de goles (a favor o en contra) son Santa Fe, Once Caldas e Independiente Medellín. Estos equipos son clave para los espectadores que buscan partidos con muchos goles. Respecto a la centralidad de intermediación, Once Caldas destaca como el equipo más influyente en las últimas temporadas, principalmente por los goles recibidos. Su rendimiento puede afectar directamente la clasificación de otros equipos, especialmente en situaciones donde la diferencia de goles es determinante. Por último Los equipos analizados mantienen una alta centralidad propia a lo largo del tiempo. Esto tiene sentido ya que son los más ganadores del fútbol colombiano, jugando más partidos y compitiendo por todos los títulos, lo que los convierte en los más influyentes en la generación de goles.



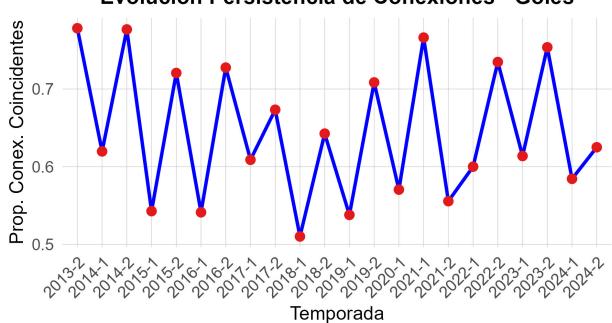
**Figura 3-12:** Evolución de la centralidad propia en la red de Goles en el periodo 2013-1 a 2024-2



**Figura 3-13:** Evolución de las comunidades en la red de Goles en el periodo 2013-1 a 2024-2



**Figura 3-14:** Evolución de la densidad en la red de Goles en el periodo 2013-1 a 2024-2



**Figura 3-15:** Evolución de la persistencia en la red de Goles en el periodo 2013-1 a 2024-2

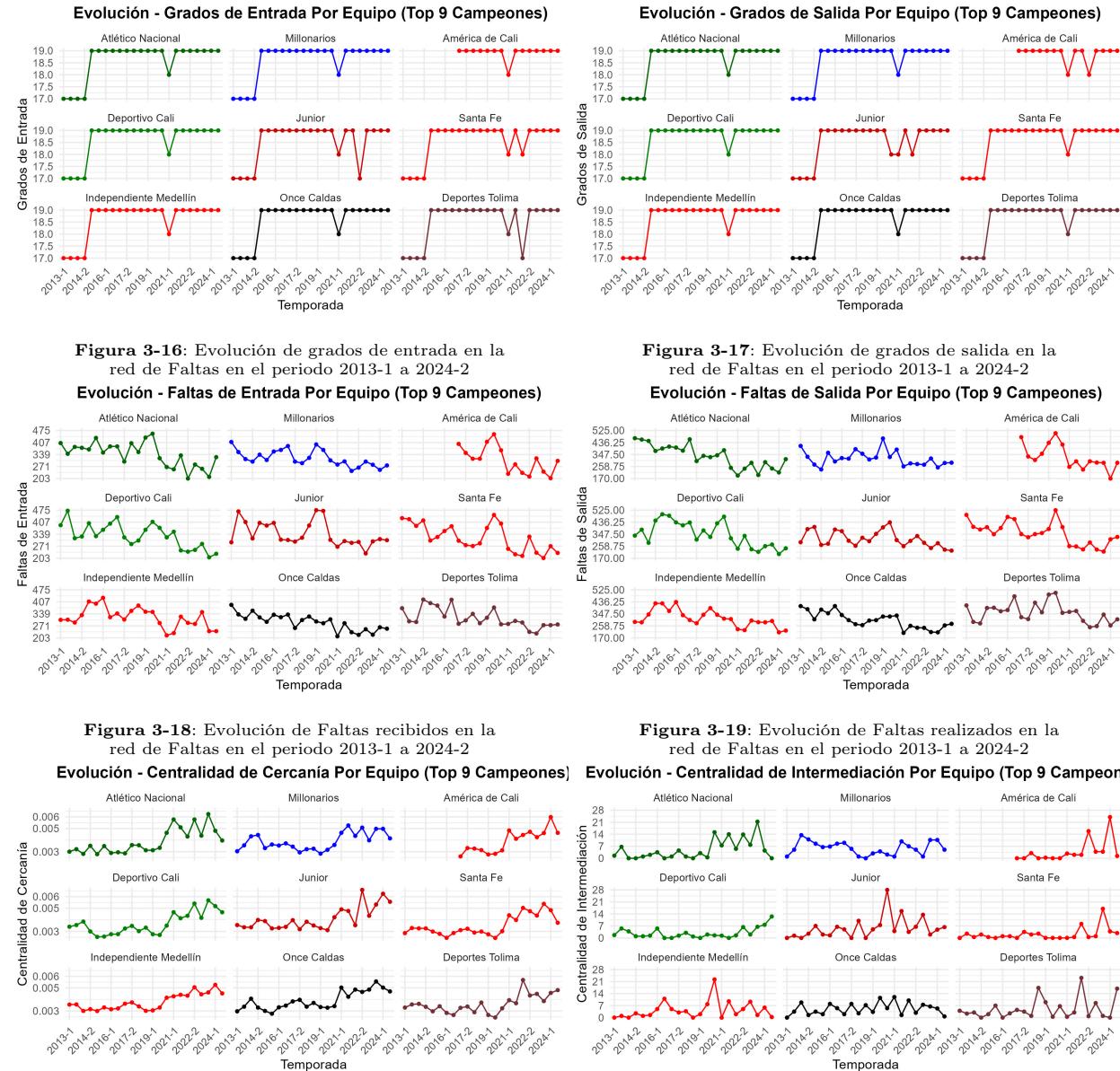
En el análisis de comunidades, se observa una primera comunidad que incluye a casi todos los equipos en todas las temporadas, y una segunda comunidad menos densa. Las comunidades 3 y 4 muestran grupos específicos como Millonarios, Deportivo Cali y Once Caldas en 2013-2; Santa Fe, Independiente Medellín y Once Caldas en 2019-1; Santa Fe y Once Caldas en 2021-2; América de Cali e Independiente Medellín en 2024-1; y Atlético Nacional y Deportivo Cali en 2024-1. Estos grupos indican dinámicas diferentes en la generación de goles comparado con los demás equipos.

La densidad de la red de goles refleja la irregularidad de la liga, especialmente en los últimos años, donde la baja densidad sugiere una disminución en la competitividad. Los años 2013, 2014 y 2016 muestran una mayor densidad, coincidiendo con una época de mayor éxito de los equipos colombianos en competencias internacionales.

La persistencia de conexiones (goles generados entre equipos) no muestra una tendencia clara, debido a la irregularidad de los equipos y al sistema de ascensos y descensos, que impide la repetición exacta de conexiones entre temporadas.

### 3.2.2 Faltas

### 3.2. Análisis Longitudinal



En los gráficos de grados de entrada y salida, lo más común es que todos los equipos alcancen el número máximo de conexiones posibles. Esto se debe a que, aunque existe la posibilidad de que un equipo no cometa faltas en un partido, lo normal es que registre al menos una. En las primeras cuatro temporadas, se observa una cantidad de conexiones más baja en todos los equipos, lo cual se explica porque en esos años la liga se jugaba con 18 equipos, a diferencia de los 20 actuales. Además, todos los equipos tienen una conexión menos en la temporada 2021-1, debido a que, por la pandemia de COVID-19, no se jugó la temporada 2020-2 y no hubo ascensos ni descensos. A esto se sumó la desafiliación del Cúcuta Deportivo por sus deudas y malos manejos financieros, lo que redujo el número de equipos a 19 en la temporada 2021-1. También hay casos aislados en los que algunos equipos tienen una o dos conexiones menos en temporadas específicas, lo cual puede deberse a partidos que no se jugaron y se decidieron por walkover (W) o a sanciones que otorgaron los puntos al rival, eliminando registros de faltas o tarjetas.

Atlético Nacional y Millonarios son los equipos que más faltas han recibido a lo largo del tiempo. Esto

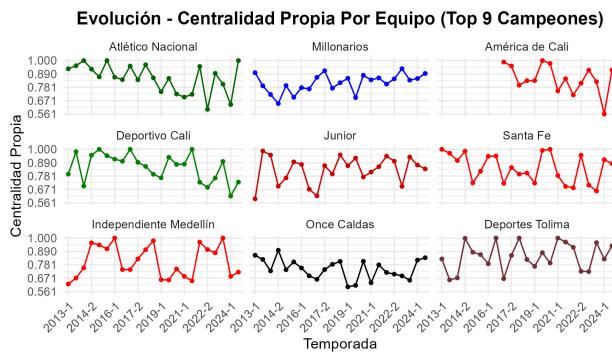
### 3. Aplicación

### 3. Aplicación

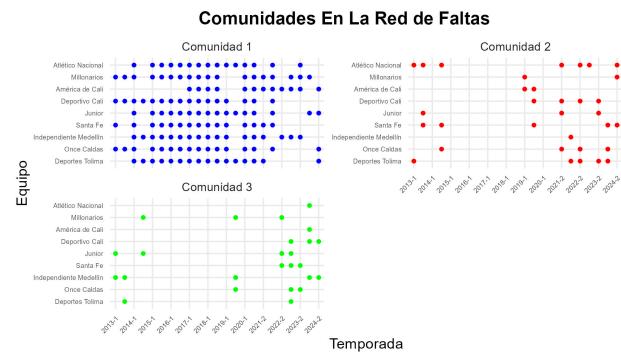
### 3.2. Análisis Longitudinal

se explica porque, al ser equipos grandes y propositivos, los rivales buscan contrarrestarlos cometiendo faltas, como la famosa falta táctica para detener contraataques. Por otro lado, Once Caldas e Independiente Medellín son los equipos que menos faltas han cometido, mostrando un estilo de juego más limpio. Además, se observa una disminución notable en las faltas recibidas y cometidas por parte de todos los equipos, lo que refleja una mejora en el juego limpio de la liga con el paso del tiempo.

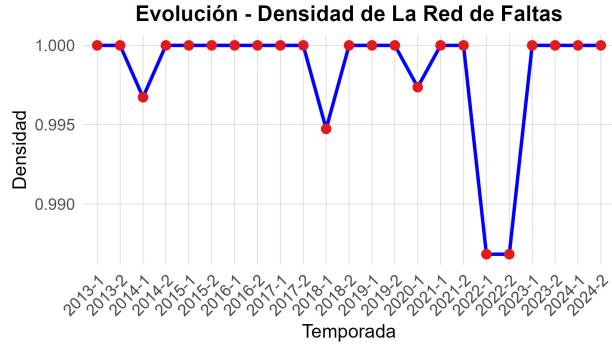
En cuanto a la centralidad de cercanía, los valores son muy bajos, lo cual es coherente, ya que el juego brusco está más asociado a equipos pequeños con menos herramientas para competir. De igual manera, la centralidad de intermediación también es baja, ya que estos equipos no son influyentes en la generación de faltas. Sin embargo, tienen una centralidad propia alta, al estar conectados con equipos influyentes (generadores de faltas), que suelen ser los equipos más pequeños.



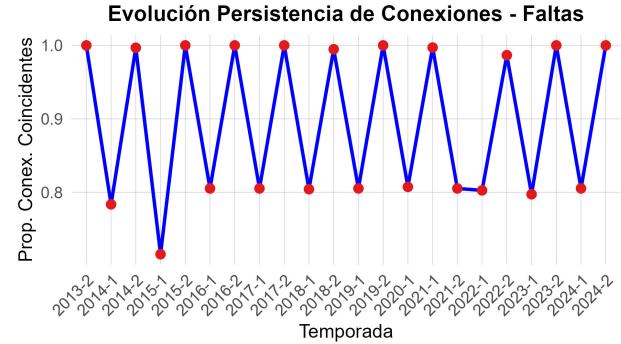
**Figura 3-22:** Evolución de la centralidad propia en la red de Faltas en el periodo 2013-1 a 2024-2



**Figura 3-23:** Evolución de las comunidades en la red de Faltas en el periodo 2013-1 a 2024-2



**Figura 3-24:** Evolución de la densidad en la red de Faltas en el periodo 2013-1 a 2024-2



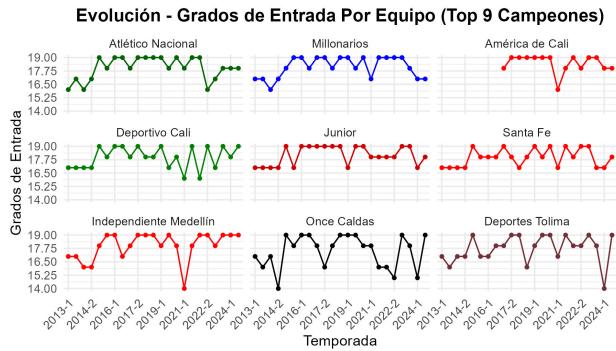
**Figura 3-25:** Evolución de la persistencia en la red de Faltas en el periodo 2013-1 a 2024-2

En el análisis de comunidades, se observa una primera comunidad muy densa que concentra a casi todos los equipos, y otras dos comunidades con menor densidad pero con una cantidad similar de equipos en distintas temporadas. Esto indica que la cantidad de faltas entre los equipos de estas comunidades pudo ser significativamente más alta o más baja en comparación con los equipos de la comunidad principal.

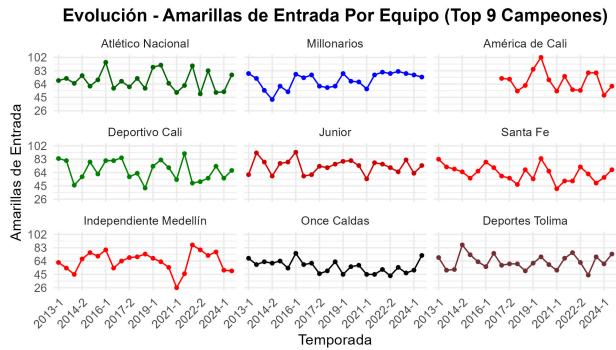
Respecto a la densidad de las redes a lo largo del tiempo, lo normal sería que siempre fuera 1. Sin embargo, como se mencionó anteriormente, hubo temporadas en las que algunos partidos se decidieron por W o por decisiones administrativas, lo que eliminó registros de faltas y tarjetas. Finalmente, la persistencia de las conexiones a lo largo del tiempo se comporta de manera similar a la red de goles, debido al sistema de ascensos y descensos, que impide que todas las conexiones se repitan entre temporadas de distintos años. En las temporadas donde no hubo descensos, la persistencia es de 1, ya que lo más común es que siempre se cometan faltas entre los equipos.

### 3.2. Análisis Longitudinal

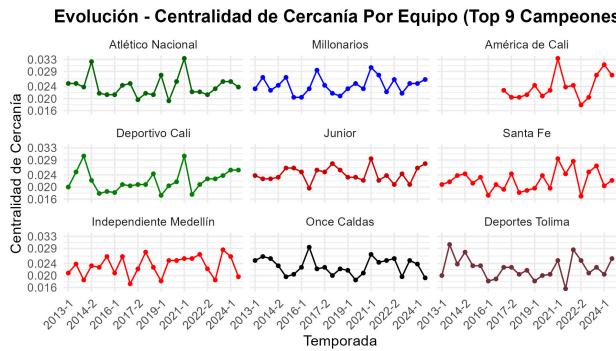
#### 3.2.3 Amarillas



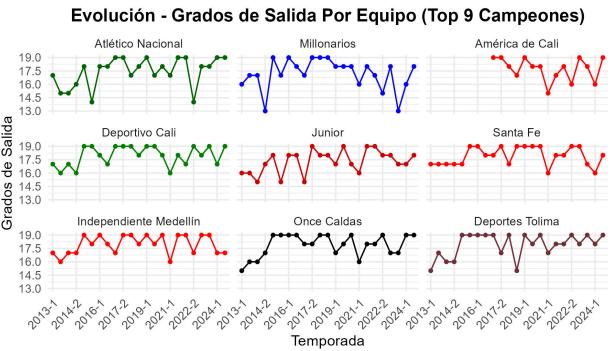
**Figura 3-26:** Evolución de grados de entrada en la red de Amarillas en el período 2013-1 a 2024-2



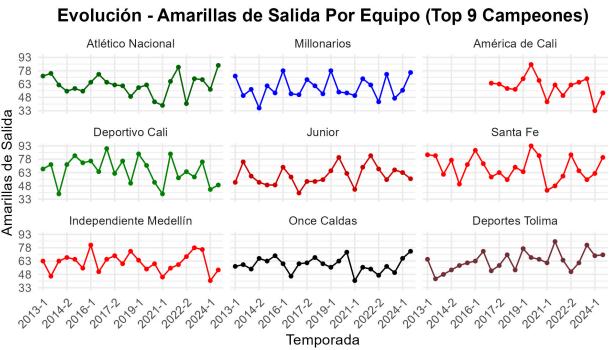
**Figura 3-28:** Evolución de Amarillas recibidos en la red de Amarillas en el período 2013-1 a 2024-2



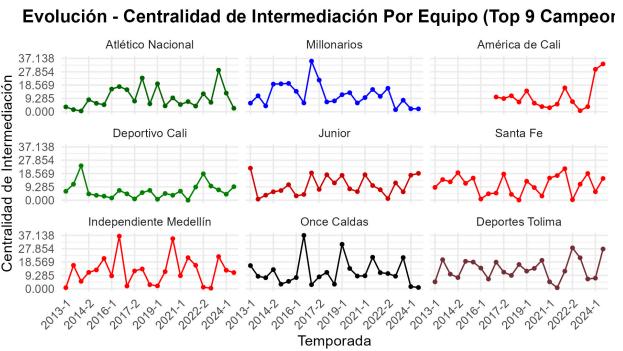
**Figura 3-30:** Evolución de la centralidad de cercanía en la red de Amarillas en el período 2013-1 a 2024-2



**Figura 3-27:** Evolución de grados de salida en la red de Amarillas en el período 2013-1 a 2024-2



**Figura 3-29:** Evolución de Amarillas realizados en la red de Amarillas en el período 2013-1 a 2024-2



**Figura 3-31:** Evolución de la centralidad de intermediación en la red de Amarillas en el período 2013-1 a 2024-2

Al igual que con las faltas, el comportamiento más común de los equipos durante los partidos es recibir al menos una tarjeta amarilla, aunque existe una mayor posibilidad de terminar un partido sin recibirla. Por ello, solo en temporadas muy específicas se observa un número menor de rivales a los que se les provocó al menos una amarilla. Por ejemplo, Independiente Medellín en 2021-1, Once Caldas en 2014-2 y Deportes Tolima en 2024-1 provocaron al menos una amarilla a solo 14 rivales. En cuanto a las amarillas recibidas, destaca Millonarios en 2014-2 y 2023-2, recibiendo al menos una amarilla contra solo 13 rivales.

Por otro lado, la mayor cantidad de amarillas generadas a sus rivales corresponde a América de Cali en 2019-2, con un total de 102. En contraste, la menor cantidad de amarillas generadas a los rivales fue registrada

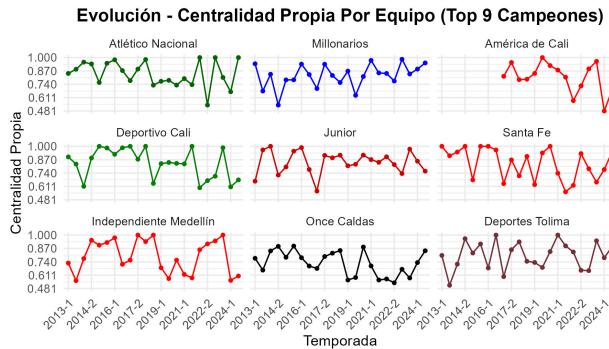
### 3. Aplicación

### 3. Aplicación

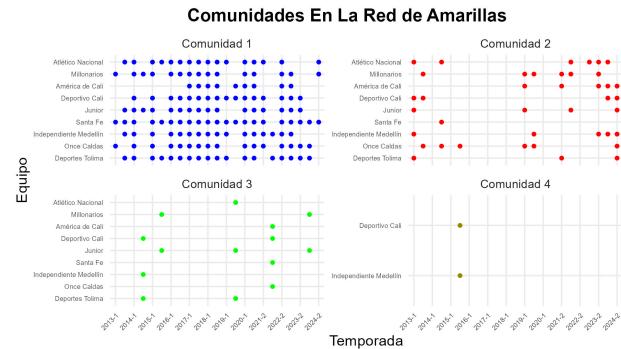
### 3.2. Análisis Longitudinal

por Independiente Medellín en 2021-1, con solo 26. En cuanto a las amarillas recibidas, Santa Fe lidera este aspecto en la temporada 2019-2, con 93 amarillas, mientras que América de Cali registró la menor cantidad en 2024-1, con solo 33.

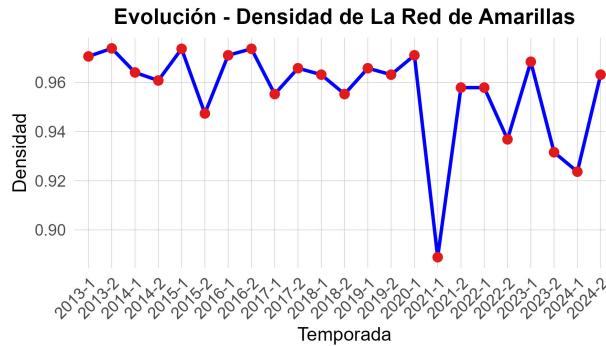
En cuanto a la centralidad de cercanía, los valores son muy bajos al igual que en la red de faltas, lo cual es coherente, ya que el juego brusco está más asociado a equipos pequeños con menos herramientas para competir. De igual manera, la centralidad de intermediación también es baja, ya que estos equipos no son influyentes en la generación de amarillas. Sin embargo, tienen una centralidad propia alta, al estar conectados con equipos influyentes (generadores de amarillas), que suelen ser los equipos más pequeños.



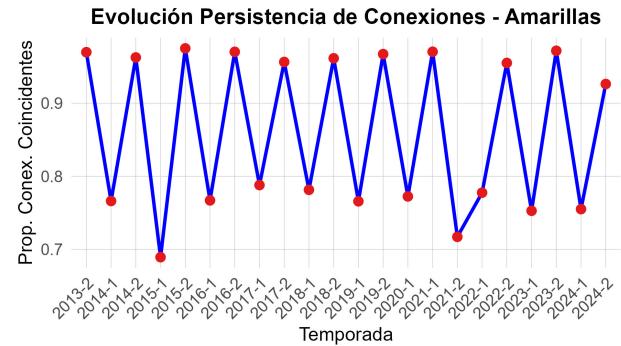
**Figura 3-32:** Evolución de la centralidad propia en la red de Amarillas en el periodo 2013-1 a 2024-2



**Figura 3-33:** Evolución de las comunidades en la red de Amarillas en el periodo 2013-1 a 2024-2



**Figura 3-34:** Evolución de la densidad en la red de Amarillas en el periodo 2013-1 a 2024-2



**Figura 3-35:** Evolución de la persistencia en la red de Amarillas en el periodo 2013-1 a 2024-2

En el análisis de comunidades, se observa una primera comunidad muy densa que concentra a casi todos los equipos, y una segunda comunidad con menor densidad pero con una cantidad considerable de equipos en distintas temporadas. Además, hay dos comunidades adicionales con grupos destacables: la comunidad 3, que incluye a Millonarios y Junior en 2015-2 y 2024-2, y la comunidad 4, que agrupa a Deportivo Cali e Independiente Medellín en 2015-2. Esto sugiere que la cantidad de tarjetas amarillas entre los equipos de estas dos últimas comunidades pudo ser significativamente más alta o más baja en comparación con los equipos de las comunidades principales.

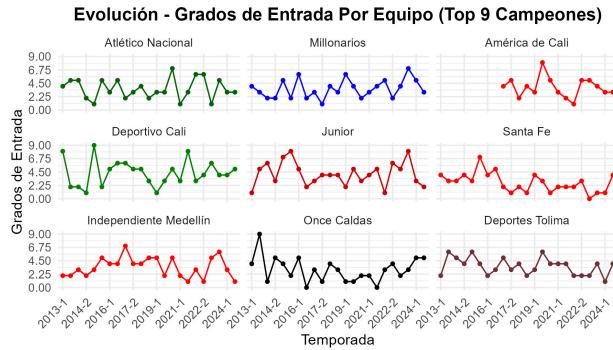
Respecto a la densidad de las redes a lo largo del tiempo, lo normal sería que esta sea alta o cercana a 1. Sin embargo, como se mencionó anteriormente, es posible que un equipo termine un partido sin recibir tarjetas amarillas. Además, en algunas temporadas, ciertos partidos se decidieron por walkover (W) o decisiones administrativas, lo que eliminó registros de faltas y tarjetas. Finalmente, la persistencia de las conexiones a lo largo del tiempo sigue un comportamiento similar al de la red de goles, debido al sistema de ascensos y descensos, que impide que todas las conexiones se repitan entre temporadas de distintos años. En las

### 3.2. Análisis Longitudinal

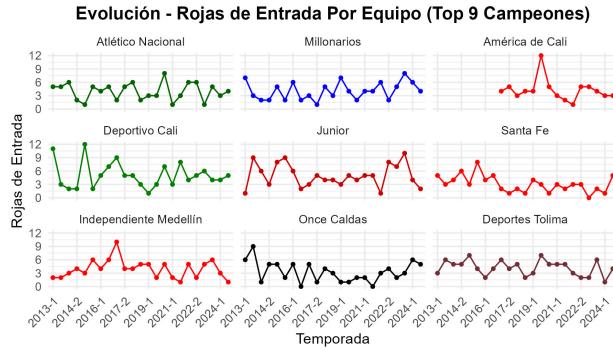
### 3. Aplicación

temporadas donde no hubo descensos, la persistencia es cercana a 1, ya que lo más común es que siempre se generen tarjetas amarillas entre los equipos.

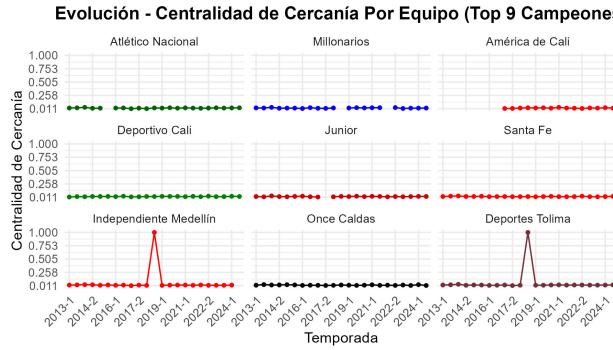
#### 3.2.4 Rojas



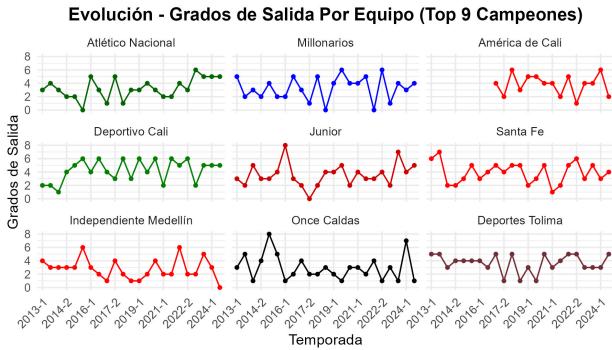
**Figura 3-36:** Evolución de grados de entrada en la red de Rojas en el periodo 2013-1 a 2024-2



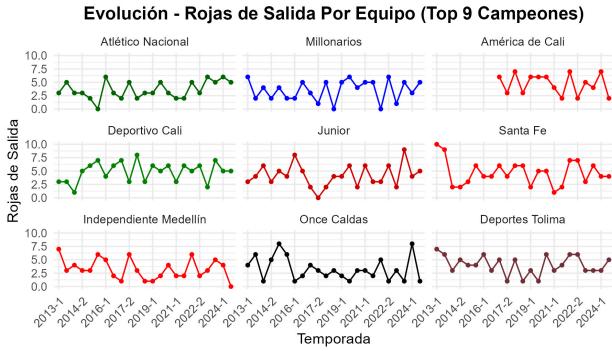
**Figura 3-38:** Evolución de Rojas recibidos en la red de Rojas en el periodo 2013-1 a 2024-2



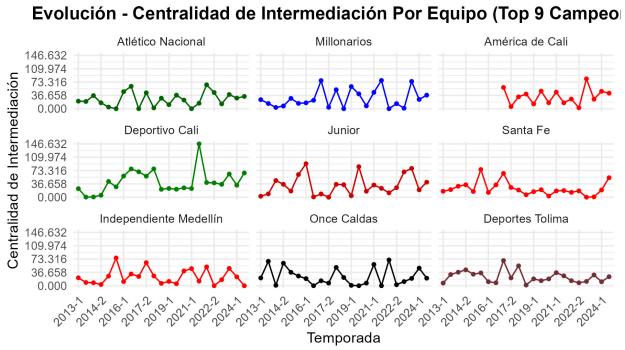
**Figura 3-40:** Evolución de la centralidad de cercanía en la red de Rojas en el periodo 2013-1 a 2024-2



**Figura 3-37:** Evolución de grados de salida en la red de Rojas en el periodo 2013-1 a 2024-2



**Figura 3-39:** Evolución de Rojas realizados en la red de Rojas en el periodo 2013-1 a 2024-2



**Figura 3-41:** Evolución de la centralidad de intermediación en la red de Rojas en el periodo 2013-1 a 2024-2

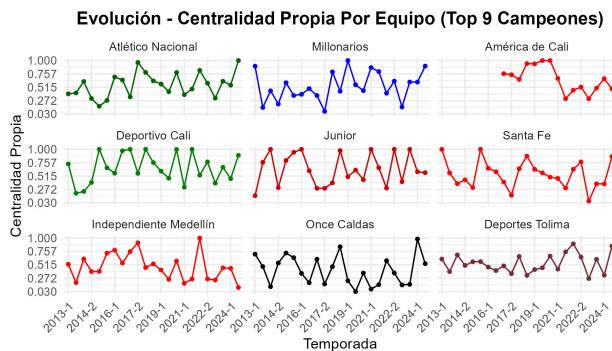
A diferencia de las faltas y las tarjetas amarillas, las tarjetas rojas no suelen ser comunes y están asociadas a un juego violento o a malos comportamientos de los jugadores dentro de la cancha. Se observa que los equipos más ganadores de Colombia tienden a generar más tarjetas rojas a sus rivales que las que reciben. Destaca Once Caldas en 2013-2, que provocó al menos una tarjeta roja a 9 rivales, mientras que Junior en 2016-1 fue el equipo contra el que más rivales recibieron expulsiones, con un total de 8.

### 3. Aplicación

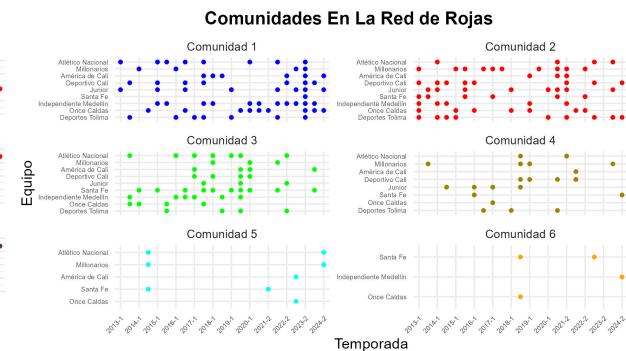
### 3.2. Análisis Longitudinal

En cuanto a las tarjetas rojas generadas, América de Cali en 2019-2 y Deportivo Cali en 2015-1 lideran con 12 rojas provocadas a sus rivales. Por otro lado, Santa Fe en 2013-1 fue el equipo con más tarjetas rojas recibidas, registrando 10, una cifra aún más destacable considerando que en esa época se jugaban menos partidos que en la actualidad.

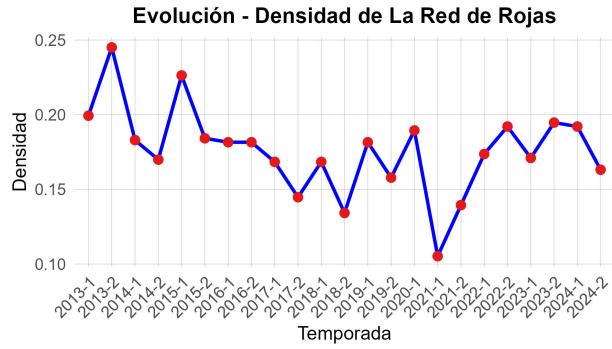
La centralidad de cercanía de los equipos más ganadores es casi nula, con excepción de Independiente Medellín y Deportes Tolima en 2018-2, donde fueron los equipos con más tarjetas rojas generadas, tanto a favor como en contra. Esto se debe a que el juego brusco suele asociarse más a equipos pequeños. La centralidad de intermediación es muy variable, sin una tendencia clara, al igual que la centralidad propia, aunque en varias temporadas Deportivo Cali y Junior alcanzaron el valor máximo posible, conectados a los equipos más generadores de rojas (equipos pequeños) en esos períodos.



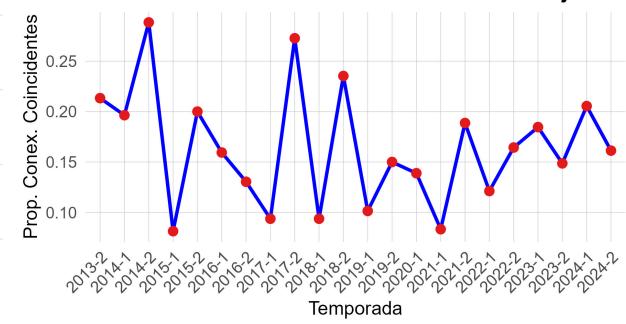
**Figura 3-42:** Evolución de la centralidad propia en la red de Rojas en el periodo 2013-1 a 2024-2



**Figura 3-43:** Evolución de las comunidades en la red de Rojas en el periodo 2013-1 a 2024-2



**Figura 3-44:** Evolución de la densidad en la red de Rojas en el periodo 2013-1 a 2024-2



**Figura 3-45:** Evolución de la persistencia en la red de Rojas en el periodo 2013-1 a 2024-2

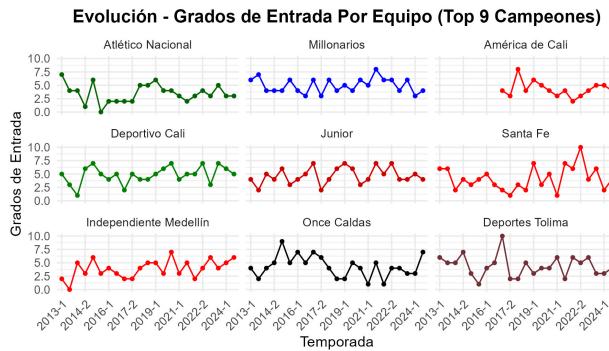
En el análisis de comunidades, se observan tres primeras comunidades con densidades similares y una cuarta con menor densidad. Destacan las comunidades 5 y 6, donde se encuentran grupos como Atlético Nacional, Millonarios y Santa Fe en 2014-2; América de Cali y Once Caldas en 2023-1; y Atlético Nacional y Millonarios en 2024-2 en la comunidad 5; y Santa Fe y Once Caldas en 2018-2 en la comunidad 6. Esto indica que la dinámica de tarjetas rojas en estas comunidades fue significativamente más alta o más baja que en las primeras cuatro.

La densidad de la red a lo largo del tiempo tiende a ser baja, ya que las tarjetas rojas no son comunes en los partidos. La temporada 2021-1 destaca como la que menos tarjetas rojas registró. En cuanto a la persistencia, los valores también son bajos, debido a la rareza de las tarjetas rojas y al sistema de ascensos y descensos. Sin embargo, la temporada 2014-2 muestra una mayor repetición de conexiones en comparación con su temporada anterior.

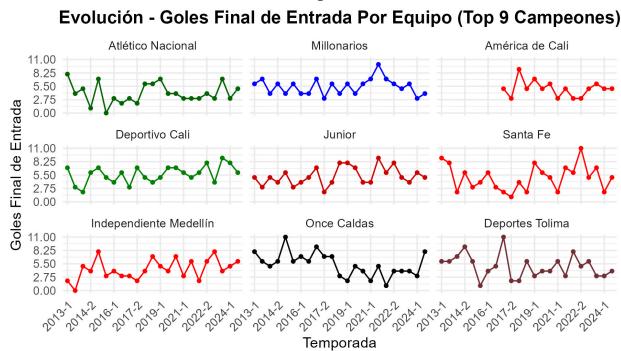
### 3.2. Análisis Longitudinal

### 3. Aplicación

#### 3.2.5 Goles Final



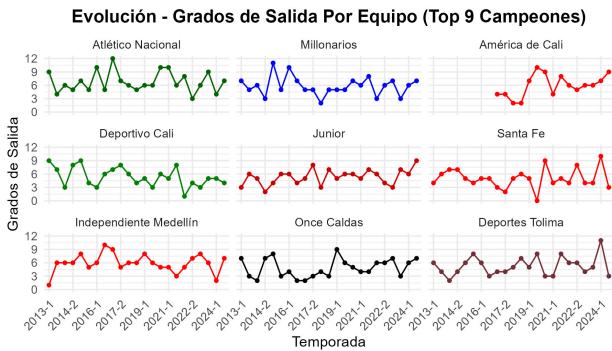
**Figura 3-46:** Evolución de grados de entrada en la red de Goles Final en el período 2013-1 a 2024-2



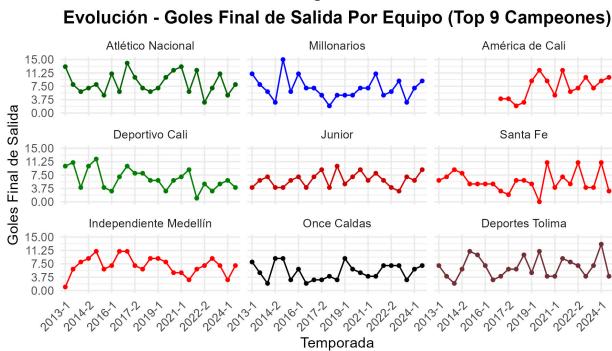
**Figura 3-48:** Evolución de Goles Final recibidos en la red de Goles Final en el período 2013-1 a 2024-2



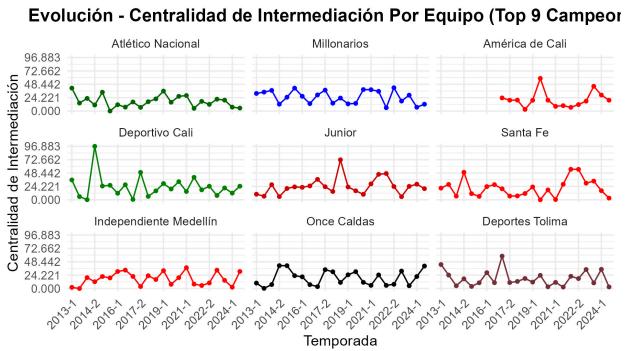
**Figura 3-50:** Evolución de la centralidad de cercanía en la red de Goles Final en el período 2013-1 a 2024-2



**Figura 3-47:** Evolución de grados de salida en la red de Goles Final en el período 2013-1 a 2024-2



**Figura 3-49:** Evolución de Goles Final realizados en la red de Goles Final en el período 2013-1 a 2024-2



**Figura 3-51:** Evolución de la centralidad de intermediación en la red de Goles Final en el período 2013-1 a 2024-2

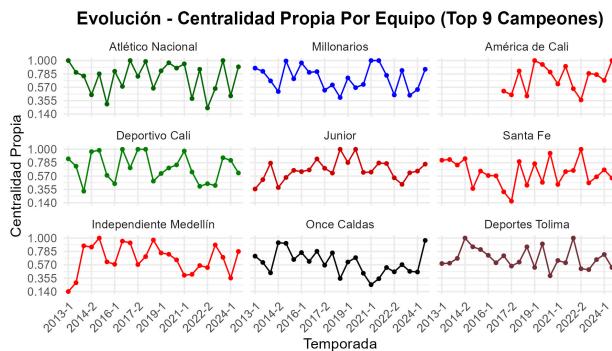
Se observa una variabilidad marcada en los grados de entrada (cantidad de equipos que anotaron al menos un gol en los últimos 15 minutos del partido) entre todos los equipos. Destacan Atlético Nacional en 2015-2 e Independiente Medellín en 2013-2, que no permitieron que ningún rival les anotara en los minutos finales, mientras que Santa Fe en 2022-2 recibió al menos un gol en los minutos finales de 10 rivales distintos. Por otro lado, los equipos más ganadores tienden a tener cifras más altas de rivales a los que les anotaron en los minutos finales, destacando Atlético Nacional en 2017-1, que marcó a 12 rivales distintos en los minutos finales, seguido por Millonarios en 2015-1 y Deportes Tolima en 2024-1, con 11 rivales cada uno. Esto se explica por lo que en el fútbol se conoce como jerarquía, esa capacidad de los equipos grandes para remontar, ganar y marcar goles en los minutos finales, manteniendo la concentración hasta el último momento.

### 3. Aplicación

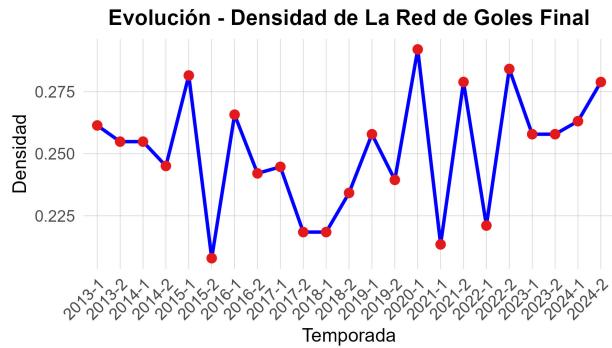
### 3.2. Análisis Longitudinal

En general, los equipos más ganadores de Colombia permiten menos goles en los minutos finales. Santa Fe en 2022-2 es la excepción, ya que permitió 11 goles en los minutos finales, mientras que Millonarios en 2015-1 fue el equipo que más goles anotó en esos momentos, con un total de 15.

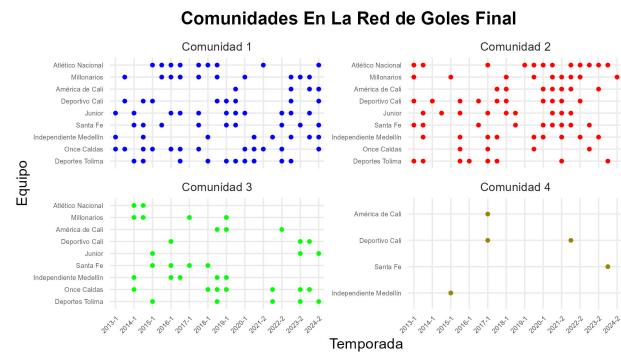
Al igual que con las tarjetas rojas, los goles en minutos finales no son comunes, por lo que la centralidad de cercanía en todos los equipos y temporadas es casi nula, excepto para Deportivo Cali en 2021-1, que fue el más influyente tanto al marcar como al recibir goles en los minutos finales. La centralidad de intermediación muestra una alta variabilidad, sin una tendencia clara. Sin embargo, equipos como Atlético Nacional, Millonarios y América de Cali tienen el valor máximo de centralidad propia en varias temporadas, ya que, aunque no son influyentes en términos generales, son los que más goles marcan en los minutos finales, conectándose entre sí y con los equipos pequeños, que suelen recibir más goles en estos momentos.



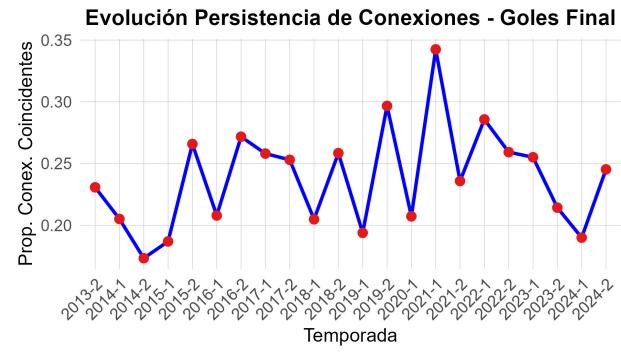
**Figura 3-52:** Evolución de la centralidad propia en la red de Goles Final en el periodo 2013-1 a 2024-2



**Figura 3-54:** Evolución de la densidad en la red de Goles Final en el periodo 2013-1 a 2024-2



**Figura 3-53:** Evolución de las comunidades en la red de Goles Final en el periodo 2013-1 a 2024-2



**Figura 3-55:** Evolución de la persistencia en la red de Goles Final en el periodo 2013-1 a 2024-2

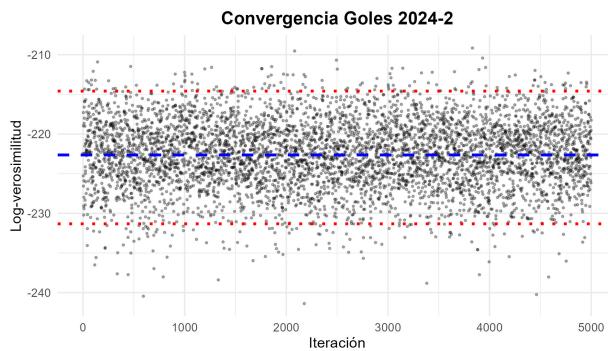
En el análisis de comunidades, se forman dos primeras comunidades grandes con densidades similares, una tercera con menor densidad y una cuarta comunidad donde destaca el grupo de América de Cali y Deportivo Cali en 2017-1. Esto sugiere que el comportamiento de goles en minutos finales en este clásico caeño fue más intenso que en otras comunidades, haciendo de este partido uno muy atractivo.

La densidad de la red tiende a ser baja, ya que los goles en minutos finales no son comunes. La temporada 2015-2 destaca por tener la menor cantidad de goles en minutos finales, mientras que la temporada 2020-1, coincidiendo con la pandemia de COVID-19, registró la mayor cantidad. La persistencia también muestra valores bajos, no solo por el sistema de ascensos y descensos, sino porque es poco común que estas conexiones se repitan en temporadas consecutivas. La temporada 2021-1 destaca por tener la mayor repetición de conexiones con respecto a la temporada 2020-1, irónicamente el único año en que no hubo ascensos ni descensos debido a la pandemia.

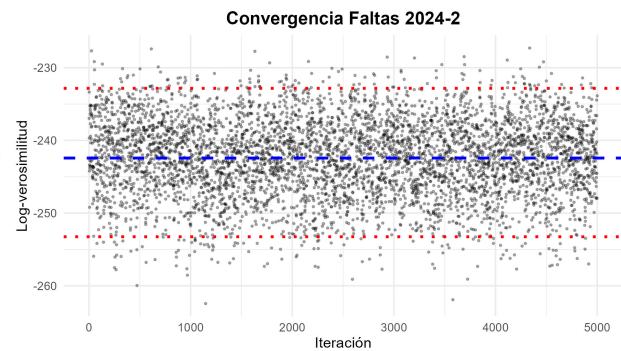
### 3.3 Modelo de espacio latente de dos dimensiones

A continuación, analizaremos el desempeño del modelo de espacio latente de dos dimensiones, el cual fue ajustado a cada una de las 115 redes construidas. Es importante destacar que, en el caso de las redes de **Faltas** y **Amarillas**, debido a su alta densidad, el modelo no lograba ajustarse adecuadamente y no convergía. Para solucionar este problema, se identificaron las aristas con mayor peso en cada una de estas redes y se estableció un umbral del 50 %, conservando solo las conexiones que superaban este valor. Este enfoque permitió entrenar el modelo de manera efectiva en estas redes.

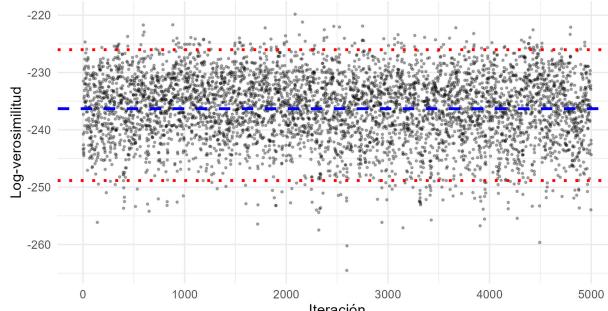
#### 3.3.1 Convergencia



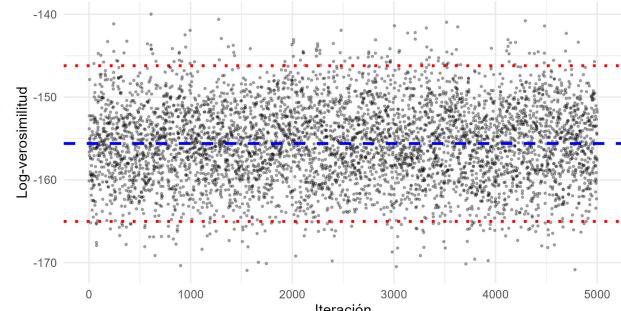
**Figura 3-56:** Convergencia Goles 2024-2  
Convergencia Amarillas 2024-2



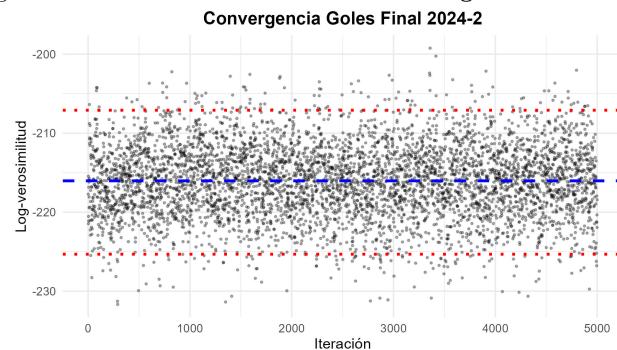
**Figura 3-57:** Convergencia Faltas 2024-2  
Convergencia Rojas 2024-2



**Figura 3-58:** Convergencia Amarillas 2024-2



**Figura 3-59:** Convergencia Rojas 2024-2



**Figura 3-60:** Convergencia Goles Final 2024-2

Como se puede observar en los gráficos de cada categoría para la temporada 2024-2, así como en los resultados incluidos en el archivo [HTML](#) mencionado anteriormente, la mayoría de los valores de la log-verosimilitud se concentran alrededor de la media y dentro del intervalo de confianza, sin mostrar patrones o tendencias claras. Esto indica que el modelo ha alcanzado un equilibrio, logrando una distribución estable sin evidencias de sobreajuste o subajuste en los datos.

### 3.3.2 Inferencia sobre el intercepto

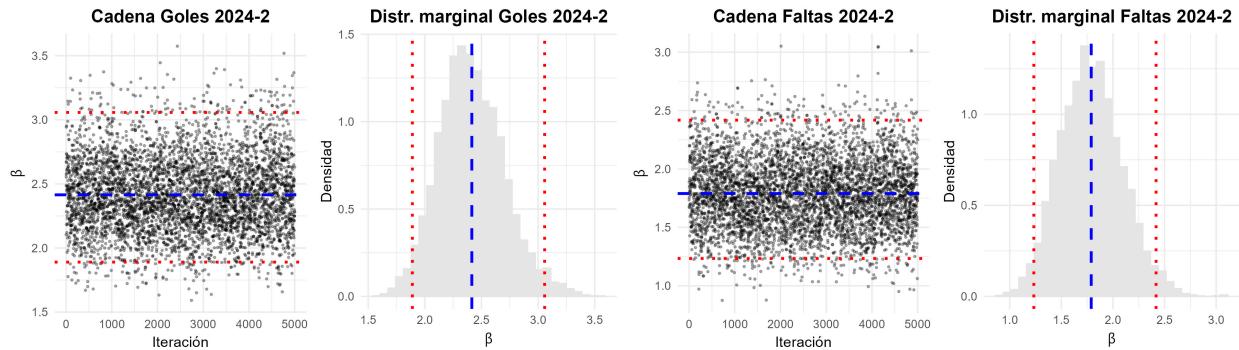


Figura 3-61: Inferencia sobre el intercepto Goles 2024-2

Figura 3-62: Inferencia sobre el intercepto Faltas 2024-2

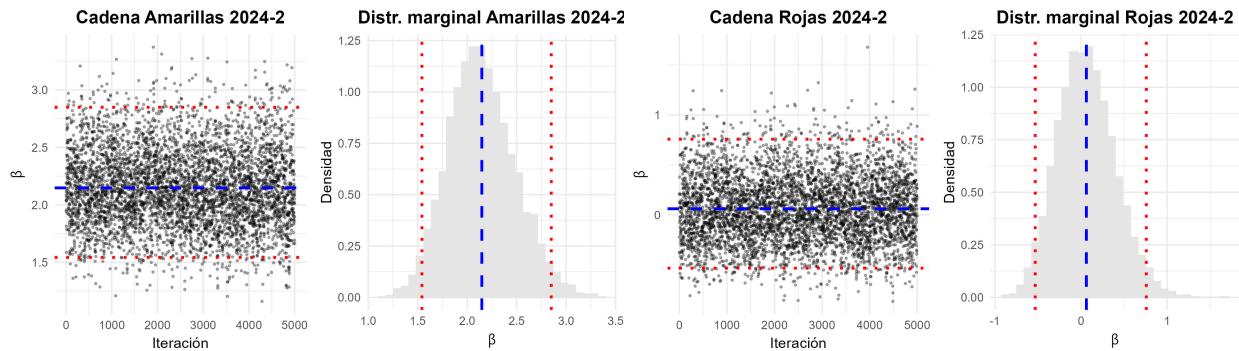


Figura 3-63: Inferencia sobre el intercepto Amarillas 2024-2

Figura 3-64: Inferencia sobre el intercepto Rojas 2024-2

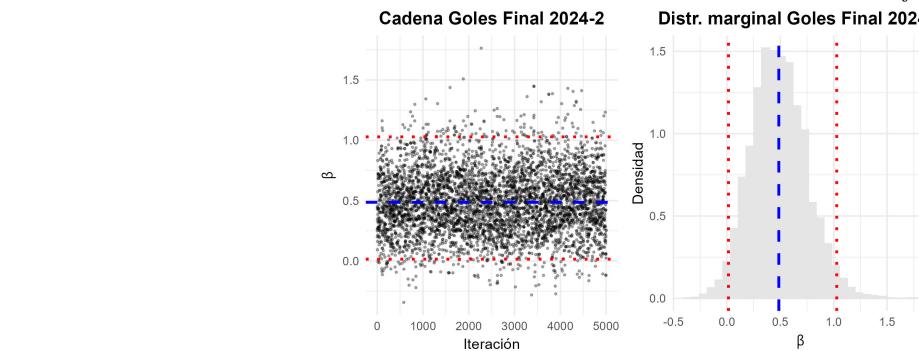


Figura 3-65: Inferencia sobre el intercepto Goles Final 2024-2

En este modelo, el intercepto ( $\beta$ ) ha convergido correctamente para todas las redes, como lo evidencia el gráfico de la cadena, donde los valores fluctúan alrededor de una media estable y dentro del intervalo de confianza,

sin mostrar tendencias o patrones no deseados. Además, el histograma de la distribución del intercepto es unimodal y suave para todas las redes, lo que indica que el modelo ha capturado adecuadamente la incertidumbre y que las estimaciones son confiables. La probabilidad de interacción basal ( $p_{\text{basal}} = \frac{1}{1+\exp(-\beta)}$ ) es siempre más cercana a 1 para todas las redes, lo que sugiere una alta probabilidad base de interacción entre los nodos (equipos) en todas las categorías analizadas. Por ejemplo, las probabilidades de interacción promedio para cada categoría son: **Goles: 0.929, Faltas: 0.877, Amarillas: 0.893, Rojas: 0.564, y Goles final: 0.602.** Estos valores reflejan una tendencia generalizada a la interacción en el modelo, lo que es coherente con el comportamiento observado en los datos. Si desea verificar, todas las gráficas de inferencia sobre el intercepto de cada una de las redes están disponibles en el archivo [HTML](#) mencionado anteriormente.

### 3.3.3 Inferencia sobre las posiciones latentes

### 3. Aplicación

### 3.3. Modelo de espacio latente de dos dimensiones

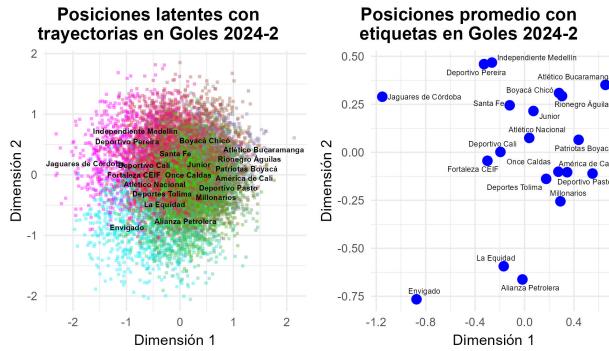


Figura 3-66: Inferencia sobre las posiciones latentes Goles 2024-2

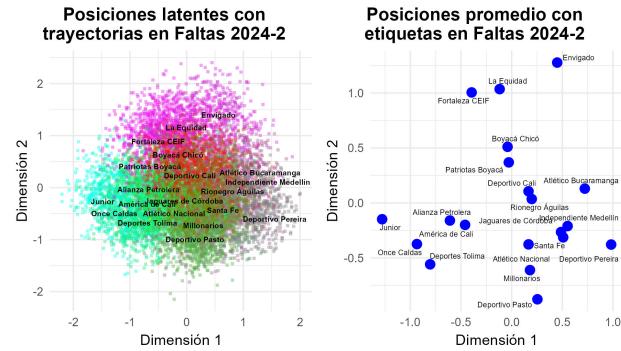


Figura 3-67: Inferencia sobre las posiciones latentes Faltas 2024-2

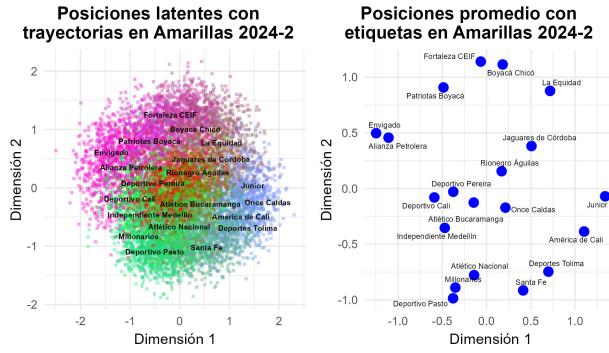


Figura 3-68: Inferencia sobre las posiciones latentes Amarillas 2024-2

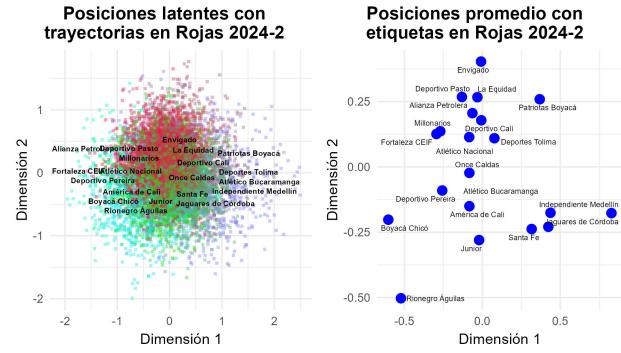


Figura 3-69: Inferencia sobre las posiciones latentes Rojas 2024-2

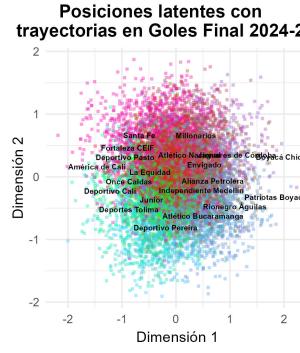


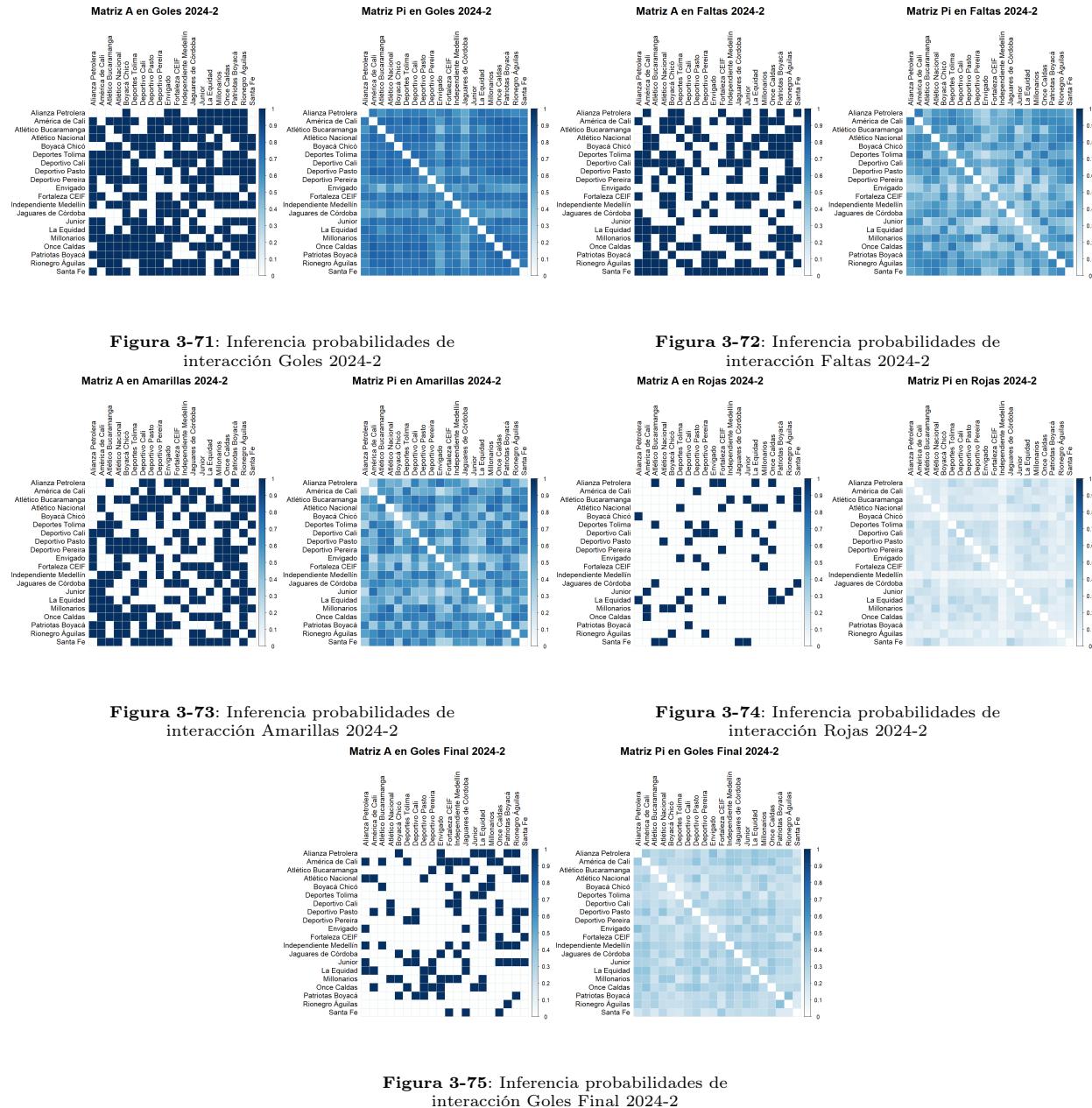
Figura 3-70: Inferencia sobre las posiciones latentes Goles Final 2024-2

En el gráfico de la izquierda, donde se muestran las trayectorias de las posiciones latentes, se observa que todos los puntos suelen concentrarse alrededor de un mismo punto para casi todas las redes (revisar el [HTML](#)), sin diferencias notorias entre los equipos, solo salvo en unos casos donde se podría alcanzar a ver la formación de 3 o 4 grupos como en la figura 3-68. Esto sugiere que, durante el proceso de muestreo, las estimaciones de las posiciones latentes presentan una alta incertidumbre y no se logra distinguir claramente patrones de interacción específicos entre los equipos. Sin embargo, en el gráfico de la derecha, que muestra las posiciones promedio, se aprecia que algunos puntos se distancian del centro, lo que indica que, al promediar las iteraciones, ciertos equipos presentan posiciones latentes diferenciadas. Esto sugiere que, aunque la incertidumbre es alta, existen equipos con patrones de interacción distintos que se destacan al analizar sus posiciones promedio. La distancia entre estos puntos refleja diferencias en sus relaciones: los equipos más cercanos tienden a interactuar de manera similar, mientras que los más alejados muestran comportamientos distintos.

### 3.3. Modelo de espacio latente de dos dimensiones

#### 3.3.4 Inferencia probabilidades de interacción

### 3. Aplicación



En este análisis, la matriz de probabilidades de interacción ( $\mathbf{P}_i$ ) muestra una distribución bastante uniforme, lo que indica que el modelo de espacio latente no distingue claramente las interacciones entre los equipos. Esto sugiere que el modelo no está capturando diferencias significativas en las relaciones entre los nodos, asignando probabilidades similares a la mayoría de las interacciones. Dado este comportamiento, es de esperar que, al comparar  $\mathbf{P}_i$  con la matriz de adyacencia original ( $\mathbf{A}$ ), no se observe una similitud notable, ya que  $\mathbf{A}$  refleja interacciones específicas y observadas, mientras que  $\mathbf{P}_i$  no logra diferenciar adecuadamente estas relaciones. Este patrón se cumple para todas las redes construidas en cada una de las 5 categorías analizadas, lo que sugiere una limitación general del modelo para representar las interacciones observadas en los datos.

### 3.3.5 Bondad de ajuste

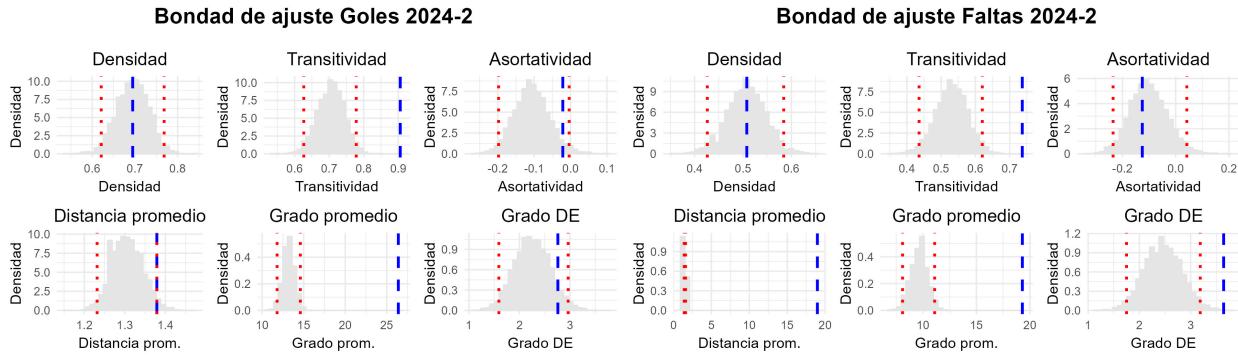


Figura 3-76: Bondad de ajuste Goles 2024-2

Figura 3-77: Bondad de ajuste Faltas 2024-2

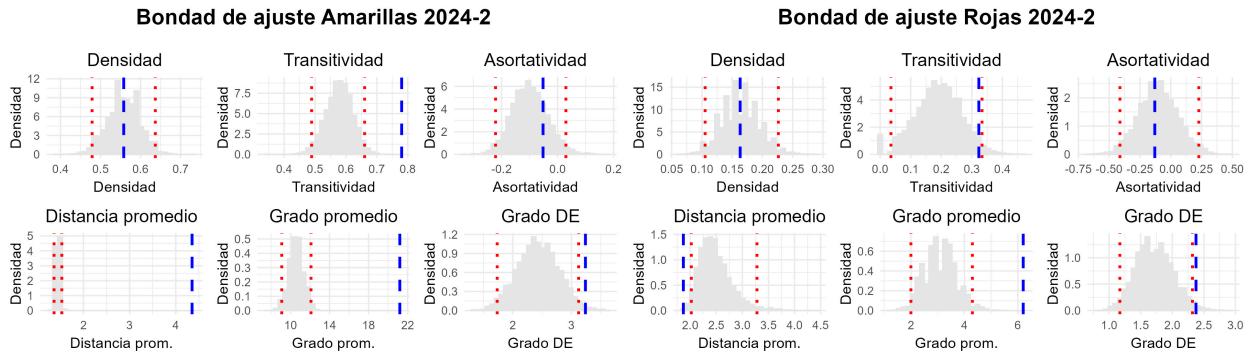


Figura 3-78: Bondad de ajuste Amarillas 2024-2

Figura 3-79: Bondad de ajuste Rojas 2024-2

### Bondad de ajuste Goles Final 2024-2

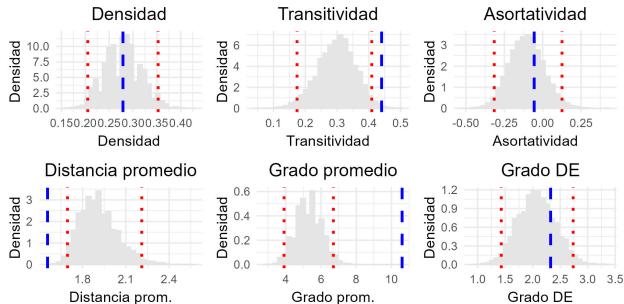


Figura 3-80: Bondad de ajuste Goles Final 2024-2

El modelo de espacio latente de dos dimensiones muestra un desempeño variable en la bondad de ajuste para las diferentes métricas evaluadas. En todas las redes, el modelo ajusta correctamente la **densidad**, lo que indica que reproduce adecuadamente el número total de conexiones observadas. La **asortatividad** también se ajusta bien en la gran mayoría de las redes, sugiriendo que el modelo captura de manera efectiva la tendencia de los nodos a conectarse con otros de características similares. Sin embargo, el modelo presenta limitaciones en otras métricas: la **distancia promedio** no se ajusta bien en la mayoría de los casos, lo que indica que no logra reproducir correctamente la estructura de caminos cortos en las redes. La **transitividad** tampoco se ajusta bien en la gran mayoría de las redes, lo que sugiere que el modelo no captura adecuadamente la formación de triángulos o agrupaciones locales. El **grado promedio** no se ajusta bien en ninguna de las redes,

lo que refleja una incapacidad del modelo para reproducir la distribución general de conexiones por nodo. Finalmente, el **grado DE** (desviación estándar del grado) se ajusta correctamente en aproximadamente el 35 % de las redes, lo que indica que el modelo solo logra capturar la variabilidad en las conexiones de los nodos en un subconjunto limitado de casos. En resumen, mientras el modelo es efectivo para ajustar la densidad y la assortatividad, presenta dificultades para reproducir otras propiedades estructurales clave de las redes.

### 3.3.6 Clustering de posiciones latentes

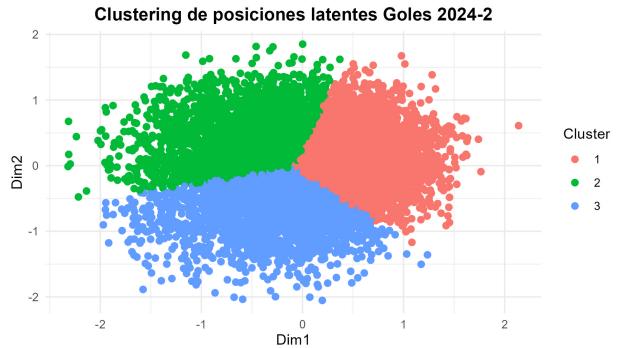
A continuación, realizaremos un análisis de **k-means** sobre las posiciones latentes con el objetivo de identificar las comunidades que emergen en función de estas posiciones.

Para determinar el número óptimo de grupos ( $k$ ), primero ajustamos un **modelo de mezcla gaussiana (GMM)** a las posiciones latentes de cada red construida. Este enfoque es válido y útil porque el modelo de mezcla gaussiana asume que los datos provienen de una combinación de varias distribuciones normales (gaussianas) y ajusta estos componentes para identificar patrones ocultos. Utilizamos el **Criterio de Información Bayesiano (BIC)** para evaluar distintos números de grupos y seleccionar el valor de  $k$  que mejor equilibra el ajuste del modelo y su complejidad. Esto ayuda a evitar el sobreajuste y proporciona una estimación más precisa del número de comunidades en los datos.

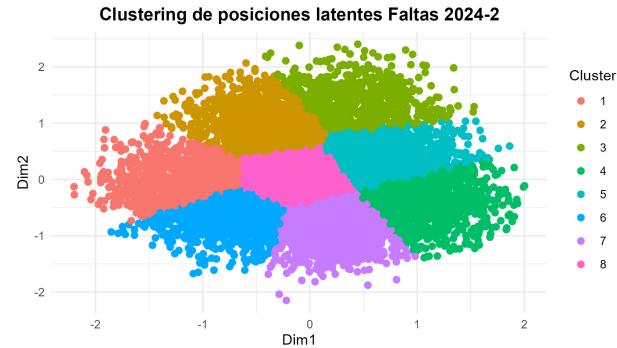
A continuación, se presentan las gráficas del resultado del *k-means* en la **temporada 2024-2** para cada una de las cinco categorías analizadas. Si desea ver el análisis completo con todas las temporadas y detalles adicionales, puede acceder al [HTML](#) mencionado anteriormente.

### 3. Aplicación

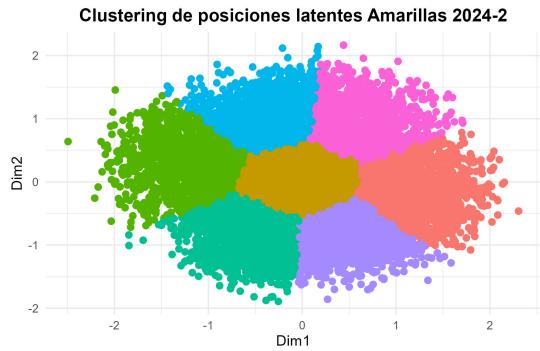
### 3.3. Modelo de espacio latente de dos dimensiones



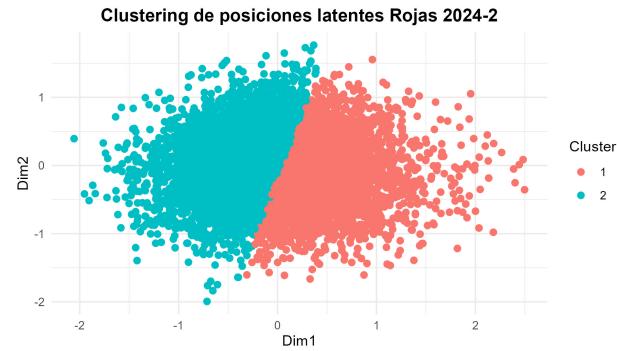
**Figura 3-81:** Clustering de posiciones latentes Goles 2024-2



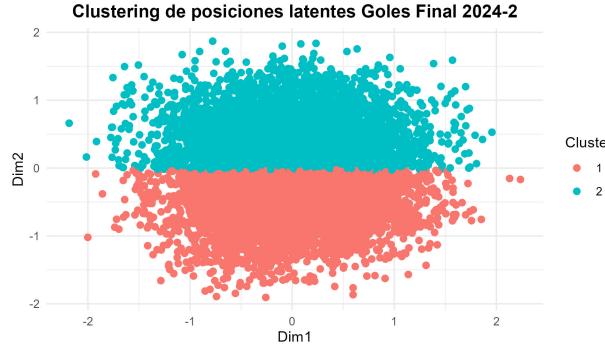
**Figura 3-82:** Clustering de posiciones latentes Faltas 2024-2



**Figura 3-83:** Clustering de posiciones latentes Amarillas 2024-2



**Figura 3-84:** Clustering de posiciones latentes Rojas 2024-2



**Figura 3-85:** Clustering de posiciones latentes Goles Final 2024-2

En el proceso de clustering aplicado a las redes de las 5 categorías, se observa que el número promedio de comunidades identificadas varía significativamente entre ellas. En la **primera categoría**, el promedio es de **3.522** comunidades, con un mínimo de **2** y un máximo de **7**. La **segunda categoría** presenta un promedio de **5.348** comunidades, con un mínimo de **3** y un máximo de **9**, lo que sugiere una mayor fragmentación en los grupos. En la **tercera categoría**, el promedio es de **5.087** comunidades, con un mínimo de **3** y un máximo de **10**, indicando una estructura aún más diversa. Por otro lado, la **cuarta categoría** muestra un promedio de **3.043** comunidades, con un mínimo de **2** y un máximo de **6**, reflejando una estructura más cohesionada. Finalmente, la **quinta categoría** tiene un promedio de **3.217** comunidades, con un mínimo de **2** y un máximo de **6**. Estos resultados resaltan la diversidad en la formación de comunidades entre las categorías, lo que refleja diferencias en los patrones de interacción entre los equipos. En la siguiente sección, analizaremos con mayor detalle las proporciones de los equipos en cada comunidad relacionada a las figuras **3-81, 3-82, 3-83, 3-84** y **3-85** de la temporada 2024-2 respectivamente, lo que permitirá identificar agrupaciones específicas y su relevancia en el contexto de cada categoría.

### 3.3.7 Proporción de los equipos en las comunidades (Cluster)

Al igual que en los gráficos de clustering basados en posiciones latentes, a continuación se presentan las tablas de proporciones correspondientes a dichos gráficos. Estas tablas detallan la distribución de los equipos en las comunidades identificadas. Si desea consultar el resto de las tablas de proporciones, puede revisar el archivo [HTML](#) mencionado anteriormente, donde se encuentran disponibles todas las tablas completas.

Red Goles 2024-2			
Equipo/Comunidades	1	2	3
Alianza Petrolera	0.023	0.015	0.127
América de Cali	0.082	0.014	0.041
Atlético Bucaramanga	0.106	0.011	0.008
Atlético Nacional	0.055	0.058	0.035
Boyacá Chicó	0.072	0.048	0.019
Deportes Tolima	0.060	0.025	0.060
Deportivo Cali	0.030	0.067	0.063
Deportivo Pasto	0.094	0.008	0.028
Deportivo Pereira	0.019	0.128	0.014
Envigado	0.001	0.030	0.143
Fortaleza CEIF	0.018	0.078	0.068
Independiente Medellín	0.026	0.117	0.014
Jaguares de Córdoba	0.000	0.144	0.025
Junior	0.051	0.068	0.030
La Equidad	0.016	0.013	0.140
Millonarios	0.062	0.015	0.069
Once Caldas	0.077	0.016	0.045
Patriotas Boyacá	0.097	0.012	0.020
Rionegro Águilas	0.077	0.042	0.018
Santa Fe	0.033	0.091	0.032

**Tabla 3-2:** Proporción de los equipos en las comunidades (Cluster) - Red Goles 2024-2

Como se observa en la tabla **3-2**, **Envigado** destaca en la comunidad 3 con una proporción de **0.143**, lo que sugiere que este equipo tiene un comportamiento único en términos de influencia en la generación de goles. Por otro lado, **Atlético Bucaramanga** tiene una mayor presencia en la comunidad 1 (**0.106**), mientras que **Jaguares de Córdoba** se concentra en la comunidad 2 (**0.144**). Estas distribuciones indican que los equipos tienden a agruparse según su influencia en los resultados ofensivos y defensivos.

Equipo/Comunidades	Red Faltas 2024-2							
	1	2	3	4	5	6	7	8
Alianza Petrolera	0.154	0.027	0.003	0.003	0.006	0.130	0.020	0.066
América de Cali	0.057	0.002	0.000	0.000	0.004	0.133	0.018	0.147
Atlético Bucaramanga	0.001	0.010	0.018	0.105	0.186	0.003	0.028	0.027
Atlético Nacional	0.000	0.000	0.000	0.036	0.025	0.023	0.138	0.096
Boyacá Chicó	0.014	0.142	0.066	0.008	0.088	0.002	0.002	0.082
Deportes Tolima	0.088	0.002	0.000	0.000	0.001	0.287	0.012	0.027
Deportivo Cali	0.012	0.033	0.012	0.036	0.112	0.018	0.033	0.101
Deportivo Pasto	0.001	0.000	0.000	0.040	0.006	0.046	0.217	0.012
Deportivo Pereira	0.000	0.002	0.005	0.283	0.057	0.002	0.034	0.006
Envigado	0.005	0.072	0.488	0.015	0.048	0.001	0.002	0.002
Fortaleza CEIF	0.035	0.283	0.103	0.002	0.018	0.004	0.001	0.015
Independiente Medellín	0.016	0.040	0.042	0.143	0.070	0.016	0.059	0.012
Jaguares de Córdoba	0.003	0.010	0.006	0.099	0.085	0.006	0.098	0.043
Junior	0.365	0.010	0.000	0.001	0.001	0.090	0.002	0.004
La Equidad	0.019	0.219	0.231	0.003	0.028	0.001	0.001	0.014
Millonarios	0.002	0.000	0.000	0.037	0.018	0.033	0.187	0.044
Once Caldas	0.204	0.010	0.000	0.000	0.000	0.187	0.014	0.021
Patriotas Boyacá	0.020	0.117	0.015	0.007	0.080	0.007	0.008	0.114
Rionegro Águilas	0.004	0.016	0.009	0.025	0.119	0.012	0.037	0.121
Santa Fe	0.001	0.004	0.000	0.157	0.051	0.003	0.091	0.045

**Tabla 3-3:** Proporción de los equipos en las comunidades (Cluster) - Red Faltas 2024-2

Como se observa en la tabla **3-3**, **Deportes Tolima** tiene una presencia dominante en la comunidad 6 (**0.287**), lo que sugiere un comportamiento particularmente influyente en la dinámica de las faltas. Por otro lado, **Envigado** se destaca en la comunidad 3 (**0.488**), indicando un patrón único en su influencia. **Junior** y **Once Caldas** también muestran una fuerte presencia en la comunidad 1 (**0.365** y **0.204**, respectivamente), lo que podría indicar similitudes en su impacto en la generación de faltas.

Equipo/Comunidades	Red Amarillas 2024-2						
	1	2	3	4	5	6	7
Alianza Petrolera	0.000	0.008	0.297	0.009	0.071	0.000	0.007
América de Cali	0.270	0.006	0.000	0.002	0.000	0.081	0.012
Atlético Bucaramanga	0.007	0.168	0.021	0.050	0.008	0.025	0.007
Atlético Nacional	0.002	0.034	0.000	0.165	0.000	0.108	0.001
Boyacá Chicó	0.011	0.010	0.015	0.002	0.151	0.001	0.217
Deportes Tolima	0.101	0.014	0.000	0.009	0.000	0.238	0.001
Deportivo Cali	0.001	0.095	0.149	0.052	0.014	0.010	0.005
Deportivo Pasto	0.004	0.008	0.009	0.222	0.000	0.063	0.000
Deportivo Pereira	0.002	0.142	0.058	0.057	0.023	0.010	0.008
Envigado	0.001	0.002	0.299	0.012	0.069	0.000	0.011
Fortaleza CEIF	0.001	0.004	0.013	0.000	0.251	0.000	0.145
Independiente Medellín	0.013	0.059	0.060	0.134	0.015	0.017	0.018
Jaguares de Córdoba	0.099	0.072	0.006	0.001	0.024	0.007	0.151
Junior	0.315	0.003	0.001	0.002	0.002	0.021	0.030
La Equidad	0.057	0.007	0.006	0.001	0.049	0.001	0.284
Millonarios	0.002	0.012	0.003	0.215	0.000	0.073	0.000
Once Caldas	0.040	0.161	0.000	0.019	0.002	0.062	0.014
Patriotas Boyacá	0.000	0.016	0.061	0.002	0.295	0.000	0.031
Rionegro Águilas	0.030	0.171	0.002	0.005	0.024	0.014	0.059
Santa Fe	0.042	0.008	0.000	0.039	0.000	0.267	0.000

**Tabla 3-4:** Proporción de los equipos en las comunidades (Cluster) - Red Amarillas 2024-2

Como se observa en la tabla 3-4, **Envigado** y **Alianza Petrolera** tienen una fuerte presencia en la comunidad 3 (**0.299** y **0.297**, respectivamente), lo que sugiere que estos equipos tienen un comportamiento similar en su influencia en la generación de tarjetas amarillas. **Fortaleza CEIF** y **La Equidad** destacan en la comunidad 5 (**0.251**) y la comunidad 7 (**0.284**), respectivamente, lo que indica patrones únicos en su impacto. **Junior** tiene una presencia significativa en la comunidad 1 (**0.315**), lo que podría reflejar un estilo de juego más influyente en este aspecto.

Red Rojas 2024-2		
Equipo/Comunidades	1	2
Alianza Petrolera	0.035	0.062
América de Cali	0.052	0.048
Atlético Bucaramanga	0.102	0.009
Atlético Nacional	0.040	0.058
Boyacá Chicó	0.019	0.074
Deportes Tolima	0.056	0.045
Deportivo Cali	0.041	0.057
Deportivo Pasto	0.030	0.065
Deportivo Pereira	0.026	0.069
Envigado	0.039	0.059
Fortaleza CEIF	0.022	0.072
Independiente Medellín	0.098	0.012
Jaguares de Córdoba	0.096	0.013
Junior	0.058	0.044
La Equidad	0.033	0.063
Millonarios	0.023	0.071
Once Caldas	0.040	0.058
Patriotas Boyacá	0.074	0.031
Rionegro Águilas	0.022	0.072
Santa Fe	0.093	0.016

**Tabla 3-5:** Proporción de los equipos en las comunidades (Cluster) - Red Rojas 2024-2

Como se observa en la tabla 3-5, **Atlético Bucaramanga**, **Independiente Medellín** y **Jaguares de Córdoba** tienen una mayor presencia en la comunidad 1 (0.102, 0.098 y 0.096, respectivamente), lo que sugiere que estos equipos tienen un comportamiento más influyente en la dinámica de las tarjetas rojas. Por otro lado, **Fortaleza CEIF** y **Rionegro Águilas** destacan en la comunidad 2 (0.072), lo que indica un impacto más limitado en este aspecto.

Red Goles Final 2024-2		
Equipo/Comunidades	1	2
Alianza Petrolera	0.045	0.054
América de Cali	0.030	0.068
Atlético Bucaramanga	0.084	0.020
Atlético Nacional	0.013	0.083
Boyacá Chicó	0.019	0.078
Deportes Tolima	0.090	0.014
Deportivo Cali	0.064	0.037
Deportivo Pasto	0.027	0.070
Deportivo Pereira	0.097	0.008
Envigado	0.040	0.059
Fortaleza CEIF	0.017	0.079
Independiente Medellín	0.055	0.045
Jaguares de Córdoba	0.028	0.070
Junior	0.091	0.013
La Equidad	0.042	0.057
Millonarios	0.008	0.087
Once Caldas	0.061	0.040
Patriotas Boyacá	0.083	0.020
Rionegro Águilas	0.087	0.016
Santa Fe	0.017	0.079

**Tabla 3-6:** Proporción de los equipos en las comunidades (Cluster) - Red Goles Final 2024-2

Como se observa en la tabla **3-6**, **Deportes Tolima**, **Junior** y **Rionegro Águilas** tienen una mayor presencia en la comunidad 1 (**0.090**, **0.091** y **0.087**, respectivamente), lo que sugiere que estos equipos son más influyentes en la dinámica de los goles en los minutos finales. Por otro lado, **Millonarios** y **Atlético Nacional** destacan en la comunidad 2 (**0.087** y **0.083**, respectivamente), lo que indica un comportamiento también influyente en este aspecto.

## 4 Discusión de resultados

La discusión de los resultados presentados en este estudio revela una serie de patrones y dinámicas interesantes en el fútbol colombiano, analizadas a través de redes construidas para diferentes categorías (goles, faltas, tarjetas amarillas, tarjetas rojas y goles finales). A continuación, se destacan los hallazgos más relevantes:

### Visualización de redes y patrones preliminares

Las redes construidas para la temporada 2024-2 muestran que los equipos más destacados, como **Atlético Nacional**, **Millonarios** y **Deportes Tolima**, tienden a tener un mayor impacto en la generación de goles, lo que se refleja en el tamaño de sus nodos. Esto sugiere que estos equipos no solo son más efectivos ofensivamente, sino que también son objetivos frecuentes de los rivales, como se observa en las conexiones de entrada. Por otro lado, equipos como **Jaguares** y **Envigado** muestran un menor impacto en la generación de goles, pero son frecuentemente superados por otros equipos, lo que indica debilidades defensivas.

En cuanto a las **faltas** y **tarjetas amarillas**, se observa que todos los equipos están altamente conectados, lo que refleja la naturaleza inherentemente física del fútbol. Sin embargo, equipos como **Santa Fe** y **Atlético Nacional** destacan por cometer más faltas, lo que podría estar relacionado con su estilo de juego más agresivo o con su capacidad para interrumpir el juego de los rivales. En el caso de las **tarjetas rojas**, aunque son menos frecuentes, se identifican equipos como **Independiente Medellín** y **Atlético Bucaramanga** que, aunque no reciben muchas expulsiones, sí las provocan en sus rivales, lo que podría indicar un juego más físico o táctico.

En la categoría de **goles final**, se destaca la capacidad de equipos como **Millonarios** y **América de Cali** para marcar goles en los últimos minutos del partido, lo que refleja su jerarquía y capacidad para mantener la concentración en momentos decisivos. Este comportamiento es crucial en partidos ajustados y puede ser un factor determinante en la obtención de puntos.

### Análisis longitudinal y evolución de las redes

El análisis longitudinal revela una alta variabilidad en el rendimiento de los equipos a lo largo del tiempo. Por ejemplo, en la categoría de **goles**, equipos como **Deportes Tolima** y **América de Cali** han tenido temporadas con defensas muy vulnerables, permitiendo que casi todos sus rivales les anotaran. Por otro lado, **Santa Fe** y **Atlético Nacional** han mostrado defensas sólidas en ciertos períodos, lo que coincide con sus éxitos en competiciones nacionales e internacionales.

#### 4. Discusión de resultados

En cuanto a las **faltas** y **tarjetas amarillas**, se observa una disminución general en su frecuencia a lo largo del tiempo, lo que podría indicar una mejora en el juego limpio o cambios en las reglas y sanciones. Sin embargo, equipos como **Atlético Nacional** y **Millonarios** siguen siendo los que más faltas reciben, lo que sugiere que son equipos que dominan el juego y obligan a sus rivales a recurrir a faltas tácticas.

En la categoría de **tarjetas rojas**, aunque son menos frecuentes, se observa que los equipos más ganadores tienden a provocar más expulsiones en sus rivales que las que reciben, lo que podría estar relacionado con su capacidad para presionar y dominar el juego. Finalmente, en la categoría de **goles final**, se destaca la consistencia de equipos como **Millonarios** y **Atlético Nacional** para marcar en los minutos decisivos, lo que refuerza su reputación como equipos con gran capacidad de reacción y mentalidad ganadora.

### Modelo de espacio latente y clustering

El modelo de espacio latente de dos dimensiones mostró un buen ajuste en términos de densidad y assortatividad, pero presentó limitaciones para capturar otras propiedades estructurales de las redes, como la transitividad y el grado promedio. Esto sugiere que, aunque el modelo es útil para identificar patrones generales de interacción, no logra capturar completamente la complejidad de las relaciones entre los equipos.

El clustering basado en las posiciones latentes reveló que el número de comunidades varía significativamente entre las categorías, lo que refleja diferencias en los patrones de interacción. Por ejemplo, en la categoría de **goles**, se identificaron comunidades claras donde equipos como **Envigado** y **Jaguares** mostraron comportamientos únicos, mientras que en la categoría de **faltas**, las comunidades estaban más dispersas, lo que sugiere una mayor uniformidad en el comportamiento de los equipos.

## 5 Conclusiones

Los resultados de este estudio ofrecen una visión detallada de las dinámicas del fútbol colombiano a través del análisis de redes, modelos de espacio latente y técnicas de clustering. A lo largo de las temporadas analizadas (2013-1 a 2024-2), se identificaron patrones de interacción, tendencias temporales y comportamientos clave que caracterizan a los equipos más relevantes de la liga. A continuación, se presentan las conclusiones más destacadas derivadas de este análisis:

- El análisis de redes permitió identificar patrones claros de interacción y conexiones entre los equipos del fútbol colombiano, destacando el impacto ofensivo y defensivo de equipos históricamente relevantes como **Atlético Nacional** y **Millonarios**. Estos equipos no solo mostraron una mayor efectividad en la generación de goles, sino que también fueron objetivos frecuentes de sus rivales, lo que refleja su centralidad en la red de interacciones y su importancia en la dinámica competitiva de la liga.
- El análisis longitudinal de las redes reveló una alta variabilidad en las conexiones y relaciones entre los equipos a lo largo del tiempo. Equipos como **Santa Fe** y **Atlético Nacional** mostraron períodos de sólidas conexiones defensivas, mientras que otros, como **América de Cali** y **Deportes Tolima**, experimentaron fases de mayor vulnerabilidad en sus redes de interacción. Además, se observó una tendencia general hacia la disminución de conexiones asociadas a faltas y tarjetas amarillas, lo que podría indicar cambios en el estilo de juego.
- El modelo de espacio latente de dos dimensiones demostró ser útil para identificar patrones generales de interacción y proximidad entre los equipos en la red. Aunque presentó limitaciones para capturar propiedades estructurales más complejas, como la transitividad y el grado promedio, el modelo permitió visualizar la proximidad latente entre equipos, lo que facilitó la identificación de comunidades mediante técnicas de clustering. Estas comunidades reflejan grupos de equipos con comportamientos similares en términos de interacciones y conexiones.
- La categoría de goles finales destacó la capacidad de equipos como **Millonarios** y **América de Cali** para marcar goles en los últimos minutos del partido, lo que refuerza su reputación como equipos con mentalidad ganadora. Este comportamiento no solo influye en los resultados de los partidos, sino que también refleja la importancia de sus conexiones y la jerarquía en momentos críticos del juego.

## 6 Recomendaciones

A partir de los hallazgos y conclusiones de este estudio, se proponen una serie de recomendaciones dirigidas a entrenadores, analistas deportivos y directivos de equipos de fútbol, con el objetivo de mejorar el rendimiento y la toma de decisiones estratégicas. Estas recomendaciones se basan en los patrones de interacción, la evolución de las redes y los comportamientos identificados en las diferentes categorías analizadas. A continuación, se presentan las principales sugerencias:

- Dado que equipos como **América de Cali** y **Deportes Tolima** mostraron períodos de vulnerabilidad en sus redes defensivas, se recomienda priorizar el entrenamiento táctico y la coordinación entre defensores. Esto incluye mejorar la comunicación en la cancha y analizar las interacciones con rivales que han demostrado ser efectivos en la generación de goles.
- Aunque se observó una disminución general en la frecuencia de faltas y tarjetas amarillas, equipos como **Atlético Nacional** y **Millonarios** siguen siendo objetivos frecuentes de faltas. Se sugiere analizar estas situaciones para identificar oportunidades de mejorar la posesión del balón y reducir la dependencia de faltas tácticas, lo que podría minimizar el riesgo de expulsiones y sanciones.
- La capacidad de equipos como **Millonarios** y **América de Cali** para marcar goles en los últimos minutos del partido es un recurso valioso. Se recomienda diseñar estrategias específicas para estos momentos críticos, como la inclusión de jugadores con alta resistencia física y mental, así como la implementación de jugadas ensayadas para maximizar la efectividad en los minutos finales.
- Dado que el modelo de espacio latente de dos dimensiones presentó limitaciones para capturar propiedades estructurales más complejas, como la transitividad y el grado promedio, se recomienda explorar el uso de modelos alternativos, como el modelo de bloques estocásticos (SBM) o un modelo de espacio latente con una dimensión mayor a dos. Estos modelos podrían ofrecer una mejor representación de las interacciones entre equipos y capturar patrones más complejos en las redes.
- Además de las categorías analizadas (goles, faltas, tarjetas amarillas, tarjetas rojas y goles finales), se recomienda realizar análisis de redes para otras estadísticas relevantes, como tiros al arco, tiros de esquina y fuera de lugar. Esto permitiría una comprensión más completa de las dinámicas del fútbol colombiano y proporcionaría insights adicionales para la toma de decisiones estratégicas.
- A pesar de sus limitaciones, el modelo de espacio latente demostró ser útil para identificar patrones de interacción y proximidad entre equipos. Se recomienda incorporar este tipo de análisis en la planificación estratégica, utilizando las posiciones latentes para identificar rivales con comportamientos similares y ajustar las tácticas de juego en función de estas similitudes.

# Referencias Bibliográficas

- Clemente, F. M.; Martins, F. M. L. & Mendes, R. S.: , 2015;** Social network analysis applied to team sports analysis; *International Journal of Performance Analysis in Sport*; **15** (3): 827--848.
- Clemente, F. M.; Andrade, G.; Praça, G. & Greco, P.: , 2016;** Network analysis in small-sided soccer games: the influence of additional players and playing position.
- Cotta, C.; Mora, A. M.; Merelo, J. J. & Merelo-Molina, C.: , 2013;** Network analysis of the 2010 fifa world cup champion team play; *Journal of Systems Science and Complexity*; **26** (1): 21--42.
- Fortunato, S.: , 2010;** Community detection in graphs; *Physics Reports*; **486** (3-5): 75--174.
- Freeman, L. C.: , 1979;** Centrality in social networks conceptual clarification; *Social Networks*; **1** (3): 215--239.
- Gormley, I. C. & Murphy, T. B.: , 2010;** A latent space model for rank data; *Statistical Network Analysis: Models, Issues, and New Directions*: 90--102.
- Grund, T. U.: , 2012;** Network structure and team performance: The case of english premier league soccer teams; *Social Networks*; **34** (4): 682--690.
- Gyarmati, L.; Kwak, H. & Rodriguez, P.: , 2014;** Searching for a unique style in soccer; *arXiv preprint arXiv:1409.0308*.
- Handcock, M. S.; Raftery, A. E. & Tantrum, J. M.: , 2007;** Model-based clustering for social networks; *Journal of the Royal Statistical Society: Series A (Statistics in Society)*; **170** (2): 301--354.
- Hoff, P. D.: , 2008;** Modeling homophily and stochastic equivalence in symmetric relational data; *Advances in Neural Information Processing Systems*; **20**: 1--8.
- Hoff, P. D.; Raftery, A. E. & Handcock, M. S.: , 2002a;** Latent space approaches to social network analysis; *Journal of the American Statistical Association*; **97** (460): 1090--1098.
- Hoff, P. D.; Raftery, A. E. & Handcock, M. S.: , 2002b;** Latent space approaches to social network analysis; *Journal of the American Statistical Association*; **97** (460): 1090--1098.
- Krivitsky, P. N. & Handcock, M. S.: , 2009;** Fitting position latent cluster models for social networks with latentnet; *Journal of Statistical Software*; **24** (5): 1--23.
- Livingston, J. A.: , 2012;** The hot hand and the cold hand in professional golf; *Journal of Economic Behavior & Organization*; **81** (1): 172--184.
- Lucey, P.; Oliver, D.; Carr, P.; Roth, J. & Matthews, I.: , 2013;** Quality vs quantity: Improved shot prediction in soccer using strategic features from spatiotemporal data; *MIT Sloan Sports Analytics Conference*.
- Newman, M. E. J.: , 2010;** *Networks: An Introduction*; Oxford University Press.
- Petersen-Wagner, R.: , 2019;** The fifa women's world cup and social network analysis.
- Pons, P. & Latapy, M.: , 2005;** Computing communities in large networks using random walks; *Journal of Graph Algorithms and Applications*; **10** (2): 191--218.