



Facultad de Ciencias Económicas y Empresariales
ICADE

FÚTBOL Y VIOLENCIA DE GÉNERO EN ANDALUCÍA Y LA REGIÓN DE MURCIA

Autor: Yago Bernard Maus Rumbao
Director: Riccardo Ciacci

1. Introducción

- 1.1 Contexto de la violencia de género y relevancia social del Servicio 016
- 1.2 Motivación del estudio: fútbol, emociones colectivas y llamadas al 016
- 1.3 Objetivos del trabajo y preguntas de investigación
- 1.4 Hipótesis iniciales del estudio
- 1.5 Estructura del trabajo

2. Marco teórico y revisión de la literatura

- 2.1 Violencia de género, violencia doméstica y factores desencadenantes
- 2.2 Eventos deportivos, frustración y comportamiento violento: evidencia internacional
- 2.3 Fútbol, alcohol y violencia en el ámbito doméstico
- 2.4 Evidencia empírica previa sobre resultados deportivos y violencia doméstica
- 2.5 Aportación esperada del presente estudio al caso español

3. Datos y fuentes de información

- 3.1 Servicio 016: bases de llamadas diarias y por franja horaria
- 3.2 Base futbolística: registro de goles y actividad deportiva
- 3.3 Tablas auxiliares y estructura del modelo de datos
- 3.4 Limitaciones y nivel de desagregación

4. Metodología

- 4.1 Diseño general
- 4.2 Unidad de análisis e integración de las bases
- 4.3 Construcción de variables deportivas
- 4.4 Bases finales de análisis
- 4.5 Análisis descriptivo en Power BI
- 4.6 Análisis estadístico y econométrico en RStudio
- 4.7 Fiabilidad e interpretación de resultados

5. Análisis descriptivo - Power BI

5.1 Análisis general

- 5.1.1 Evolución general de las llamadas al 016*
- 5.1.2 Diferencias por día de la semana y tipo de día*
- 5.1.3 Distribución territorial de las llamadas*
- 5.1.4 Actividad futbolística registrada por provincia*
- 5.1.5 Síntesis del análisis descriptivo general*

5.2 Análisis diario: contraste descriptivo de hipótesis

- 5.2.1 Hipótesis 1: los días con actividad futbolística registrada se asocian con un mayor número de llamadas al 016*
- 5.2.2 Hipótesis 2: las derrotas se asocian con un mayor número de llamadas*
- 5.2.3 Hipótesis 3: a mayor número de derrotas en la provincia, mayor número de llamadas*
- 5.2.4 Hipótesis 4: los partidos de tensión alta se asocian con más llamadas*
- 5.2.5 Hipótesis 5: las derrotas tardías y en descuento se asocian con más llamadas*
- 5.2.6 Hipótesis 6: las remontadas sufridas se asocian con más llamadas*
- 5.2.7 Hipótesis 7: las derrotas abultadas se asocian con más llamadas*
- 5.2.8 Hipótesis 8: los derbis se asocian con más llamadas*
- 5.2.9 Hipótesis 9: las eliminaciones se asocian con más llamadas*
- 5.2.10 Hipótesis 10: las rachas negativas previas se asocian con más llamadas*
- 5.2.11 Hipótesis 11: los goles en contra se asocian con más llamadas*
- 5.2.12 Síntesis del contraste de hipótesis diarias*

5.3 Análisis exploratorio por franjas horarias

- 5.3.1 Distribución horaria general de las llamadas*
- 5.3.2 Llamadas y partidos durante la franja*
- 5.3.3 Llamadas y franja posterior al partido*
- 5.3.4 Llamadas y goles en contra durante la franja*

5.3.5 Síntesis del análisis por franjas

5.4 Conclusión del análisis descriptivo

6. Análisis estadístico y econométrico - RStudio

6.1 Preparación de las bases para el análisis estadístico

6.1.1 Base diaria provincia-fecha

6.1.2 Base mensual provincia-franja horaria

6.1.3 Tratamiento de variables y controles temporales

6.2 Análisis exploratorio inicial

6.2.1 Estadísticos descriptivos de las variables principales

6.2.2 Matriz de correlaciones entre llamadas y variables futbolísticas

6.2.3 Comparación de medias entre días con y sin eventos deportivos

6.3 Modelos sobre la base diaria

6.4 Resultados de los modelos diarios

6.4.1 Modelos lineales con errores robustos

6.4.2 Modelos de conteo: Poisson y binomial negativa

6.4.3 Multicolinealidad y fiabilidad

6.5 Análisis exploratorio por franjas horarias

6.5.1 Estadísticos descriptivos de la base por franjas

6.5.2 Correlaciones y modelos por franjas

6.5.3 Ejercicio predictivo exploratorio

6.6 Conclusión del análisis estadístico y econométrico

7. Discusión y conclusiones

7.1 Síntesis y discusión de los resultados

7.2 Relación con la literatura previa

7.3 Implicaciones, limitaciones y líneas futuras

7.4 Conclusión final

8. Bibliografía

9. Anexos

9.1 Script de ejemplo de construcción de la base futbolística: Sevilla FC, temporada 20-21

9.2 Tablas auxiliares y ejemplos de datasets finales

9.3 Gráficos adicionales del análisis estadístico y econométrico en Rstudio

9.4 Script del análisis en RStudio

9.5 Declaración de uso de herramientas de inteligencia artificial generativa

1. Introducción

1.1 Contexto de la violencia de género y relevancia social del Servicio 016

La violencia de género constituye uno de los problemas sociales más graves y persistentes de las sociedades contemporáneas. No se trata únicamente de una cuestión penal o privada, sino de un fenómeno estructural que afecta a la seguridad, la salud, la autonomía y el bienestar de las mujeres. La Organización Mundial de la Salud define la violencia contra las mujeres como un problema de salud pública y una vulneración de los derechos humanos, señalando que alrededor de una de cada tres mujeres en el mundo ha sufrido violencia física o sexual por parte de una pareja o expareja, o violencia sexual por parte de terceros, a lo largo de su vida (OMS, 2024).

En el contexto español, la violencia de género ha sido objeto de una atención institucional creciente durante las últimas décadas. La aprobación de la Ley Orgánica 1/2004 de Medidas de Protección Integral contra la Violencia de Género supuso un punto de inflexión en la

respuesta pública, al reconocer esta violencia como una manifestación de la desigualdad entre hombres y mujeres y no como un conjunto de conflictos aislados dentro del ámbito familiar. Desde entonces, se han desarrollado instrumentos de prevención, protección, asistencia y seguimiento, entre los que destaca el Servicio 016.

El Servicio 016 es un teléfono de atención, información y asesoramiento dirigido a víctimas de violencia de género, a su entorno y a cualquier persona que conozca o sospeche de una situación de violencia. Su función es especialmente relevante porque permite una vía de contacto accesible, confidencial y no necesariamente vinculada a la interposición inmediata de una denuncia. En este sentido, las llamadas al 016 no recogen todos los episodios de violencia de género, ni pueden interpretarse como una medición directa del número real de agresiones. Sin embargo, sí constituyen un indicador útil de demanda de ayuda, información y orientación institucional.

Esta característica convierte al 016 en una fuente especialmente interesante para el análisis social y económico aplicado. Frente a otros indicadores, como denuncias policiales o sentencias judiciales, las llamadas al 016 permiten aproximarse a momentos en los que las víctimas o su entorno buscan apoyo, incluso aunque posteriormente no se active un procedimiento formal. Por tanto, estudiar la evolución temporal y territorial de estas llamadas puede aportar información sobre situaciones de tensión, visibilidad o necesidad de asistencia vinculadas a la violencia de género.

El presente trabajo parte de esta idea: las llamadas al 016 se utilizan como una aproximación indirecta a la demanda de atención en violencia de género. A partir de esta fuente, se analiza si determinados eventos futbolísticos pueden estar asociados a variaciones de corto plazo en el volumen de llamadas. El interés del estudio no reside en afirmar que el fútbol cause violencia de género de forma directa, sino en explorar si ciertos contextos deportivos, por su carga emocional y social, coinciden con aumentos en la utilización de un recurso público de atención.

1.2 Motivación del estudio: fútbol, emociones colectivas y llamadas al 016

El fútbol ocupa un lugar central en la vida social española. Más allá de su dimensión deportiva, constituye un fenómeno cultural, económico y mediático que moviliza identidades locales, rivalidades territoriales y emociones colectivas. Los partidos de fútbol pueden generar alegría, frustración, tensión, euforia o decepción, especialmente cuando se trata de encuentros de gran importancia competitiva, derbis, eliminatorias o derrotas inesperadas.

Desde el punto de vista económico y social, este tipo de eventos resulta relevante porque permite analizar cómo shocks emocionales externos pueden estar asociados a comportamientos individuales o familiares. La literatura internacional ha estudiado esta cuestión en distintos contextos, especialmente en Estados Unidos y Reino Unido. Uno de los trabajos más influyentes es el de Card y Dahl (2011), que analiza la relación entre resultados inesperados en partidos de fútbol americano y violencia doméstica. Sus resultados muestran que las derrotas inesperadas de equipos locales favoritos se asocian con un aumento de la violencia ejercida por hombres contra sus parejas en el hogar. Este trabajo es especialmente importante porque no se limita a estudiar la presencia de deporte, sino que analiza el

componente emocional del resultado: perder cuando se esperaba ganar.

En el ámbito británico también existen estudios que relacionan grandes torneos de fútbol con incrementos en incidentes de violencia doméstica. Algunos trabajos han mostrado que los partidos de la selección inglesa pueden coincidir con aumentos de denuncias o llamadas relacionadas con violencia doméstica, especialmente cuando el resultado es negativo o cuando el partido concentra una fuerte atención social. Este tipo de evidencia sugiere que los eventos deportivos pueden actuar como desencadenantes de corto plazo en contextos donde ya existen situaciones previas de tensión o violencia.

La motivación del presente TFG surge precisamente de trasladar esta pregunta al caso español. España es un país donde el fútbol tiene una enorme presencia social, con una estructura territorial muy marcada: los equipos no solo representan clubes deportivos, sino también ciudades, provincias y comunidades autónomas. Esta dimensión territorial permite plantear una pregunta empírica concreta: si un equipo vinculado a una provincia experimenta un resultado deportivo negativo o emocionalmente intenso, ¿se observa algún cambio en las llamadas al 016 de esa provincia?

El trabajo se centra en Andalucía y la Región de Murcia, territorios que presentan varios equipos con participación en competiciones nacionales durante el periodo analizado. Esta elección permite estudiar provincias con distinta intensidad futbolística, distintos volúmenes de llamadas y diferentes trayectorias deportivas. Además, el uso de datos diarios por provincia y de datos mensuales por franja horaria permite combinar una aproximación territorial y temporal, aunque siempre dentro de las limitaciones de desagregación disponibles.

La principal aportación del estudio es construir una base de datos propia que combina información del Servicio 016 con eventos futbolísticos detallados. A diferencia de un análisis que simplemente distinga entre días con partido y días sin partido, este trabajo incorpora variables más específicas: derrotas, victorias, goles en contra, diferencia de goles, remontadas sufridas, derrotas en descuento, derbis, eliminaciones, rachas negativas, partidos de tensión alta y distribución horaria de goles y partidos. De esta forma, el análisis no se limita a preguntar si “hay fútbol”, sino qué tipo de evento futbolístico se produce.

1.3 Objetivos del trabajo y preguntas de investigación

El objetivo general del trabajo es analizar si existe una asociación entre la actividad futbolística registrada y el volumen de llamadas pertinentes al Servicio 016 en Andalucía y la Región de Murcia durante el periodo 2021-2025.

Este objetivo general se concreta en varios objetivos específicos:

- se pretende construir una base de datos integrada que combine llamadas al 016 por provincia con información detallada de partidos y goles de equipos vinculados a los territorios analizados;
- se busca describir la evolución temporal y territorial de las llamadas al 016, diferenciando entre provincias, días de la semana y tipo de día;
- se analiza si los días con actividad futbolística registrada presentan un volumen medio de llamadas diferente a los días sin actividad;

- se estudia si determinados eventos deportivos de mayor intensidad emocional (como derrotas en descuento, remontadas sufridas, derbis o partidos de tensión alta) se asocian con un mayor número de llamadas;
- y, finalmente, se plantea un análisis estadístico en RStudio para contrastar si los patrones observados descriptivamente se mantienen al incorporar controles temporales y territoriales.

A partir de estos objetivos, las preguntas de investigación principales son las siguientes:

1. ¿Presentan los días con actividad futbolística registrada un volumen medio de llamadas al 016 distinto al de los días sin actividad futbolística?
2. ¿Existe una relación entre resultados deportivos negativos (derrotas, goles en contra o diferencia de goles desfavorable) y el número de llamadas al 016?
3. ¿Son los eventos deportivos emocionalmente más intensos, como derbis, eliminaciones, derrotas en descuento o remontadas sufridas, los que muestran una asociación más clara con las llamadas?
4. ¿Difieren estos patrones entre días laborables y fines de semana?
5. ¿El análisis por franjas horarias permite identificar alguna relación entre llamadas, partidos en juego, goles en contra o periodo posterior al encuentro?

Estas preguntas se formulan desde una perspectiva exploratoria y aplicada. El trabajo no pretende demostrar causalidad directa, ya que los datos disponibles no permiten observar episodios individuales ni vincular una llamada concreta con un partido específico. El objetivo es estudiar asociaciones estadísticas y patrones descriptivos que puedan contribuir a la evidencia existente sobre eventos deportivos y violencia de género.

1.4 Hipótesis iniciales del estudio

A partir de la literatura revisada y de la lógica del fenómeno analizado, el trabajo plantea una serie de hipótesis iniciales.

1ª Hipótesis: Los días con actividad futbolística registrada presentan, en promedio, un mayor número de llamadas al 016 que los días sin actividad futbolística. Esta hipótesis parte de la idea de que el fútbol puede actuar como un evento social relevante, capaz de alterar rutinas, convivencia, consumo de ocio y estados emocionales.

2ª Hipótesis: Las derrotas se asocian con un mayor número de llamadas que las victorias o los empates. Esta hipótesis se inspira en la literatura que vincula las derrotas inesperadas o emocionalmente negativas con aumentos de violencia en el hogar. No obstante, el trabajo reconoce desde el inicio que no todas las derrotas tienen la misma intensidad emocional.

3ª Hipótesis: Los eventos deportivos de alta tensión presentan una asociación más clara con las llamadas que los resultados deportivos generales. Dentro de esta categoría se incluyen derbis, eliminaciones, partidos en las últimas jornadas, remontadas sufridas, derrotas en descuento y derrotas en partidos de tensión alta. La lógica es que estos eventos pueden generar mayor frustración, rivalidad o implicación emocional que un partido ordinario.

4ª Hipótesis: La relación entre fútbol y llamadas puede variar entre días laborables y fines de semana. Esta distinción es importante porque las llamadas al 016 y la actividad futbolística presentan patrones semanales distintos. Los partidos se concentran especialmente en fines de semana, mientras que las llamadas pueden tener mayor intensidad en días laborables. Por tanto, analizar ambos tipos de día de forma conjunta podría llevar a interpretaciones confusas.

5ª Hipótesis: Se refiere al análisis horario. Se plantea que, si existiera una relación temporal de corto plazo, esta podría aparecer durante el partido, en el momento de los goles o en la franja posterior al encuentro. Sin embargo, dado que los datos de llamadas por franja horaria están agregados a nivel mensual, esta hipótesis se analiza únicamente de forma exploratoria.

Estas son las hipótesis principales del trabajo, aunque dentro de cada una se plantearán otras hipótesis más específicas, que abarcaremos en el apartado 5 del documento.

1.5 Estructura del Trabajo

El trabajo se organiza en nueve apartados. Tras esta introducción, el segundo apartado presenta el marco teórico y la revisión de la literatura. En él se revisan las principales aportaciones sobre violencia de género, factores desencadenantes, eventos deportivos, alcohol y violencia doméstica, con especial atención a los estudios que han analizado la relación entre resultados deportivos y violencia en el hogar.

El tercer apartado describe los datos y fuentes de información utilizados. Se explican las bases de llamadas al 016, la construcción de la base futbolística, las tablas auxiliares y las principales limitaciones derivadas del nivel de desagregación disponible.

El cuarto apartado desarrolla la metodología. Se presenta la unidad de análisis, la integración de las bases, la construcción de variables deportivas y las dos bases finales utilizadas para el análisis estadístico: una diaria por provincia y otra mensual por provincia y franja horaria.

El quinto apartado recoge el análisis descriptivo realizado en Power BI. En él se analiza la evolución general de las llamadas, la actividad futbolística registrada, las diferencias entre semana y fin de semana, y las principales hipótesis descriptivas asociadas a las variables deportivas construidas.

El sexto apartado presenta el análisis estadístico y econométrico en RStudio. En esta parte se estiman modelos de regresión y modelos para datos de conteo, además de análisis exploratorios de correlación, comparación de medias y posibles ejercicios predictivos.

El séptimo apartado recoge la discusión y conclusiones del trabajo. En él se sintetizan los hallazgos principales, se comparan con la literatura existente, se presentan las limitaciones y se proponen líneas futuras de investigación.

Finalmente, el trabajo se completa con la bibliografía y los anexos, donde se incluyen scripts, datasets utilizados, tablas auxiliares y resultados complementarios.

2. Marco teórico y revisión de la literatura

2.1 Violencia de género, violencia doméstica y factores desencadenantes

La violencia de género es un fenómeno complejo que no puede explicarse únicamente por factores individuales o situaciones aisladas. La literatura la entiende como una manifestación de relaciones desiguales de poder entre hombres y mujeres, sostenida por normas sociales, estructuras culturales, desigualdad económica y patrones de control dentro de las relaciones de pareja. Desde esta perspectiva, los episodios concretos de violencia no aparecen de forma desconectada, sino dentro de un contexto previo de desigualdad, dominación o conflicto.

La Organización Mundial de la Salud señala que la violencia contra las mujeres, especialmente la violencia ejercida por la pareja o expareja, constituye un problema de salud pública de alcance global. Además de las agresiones físicas, incluye violencia psicológica, sexual, control, amenazas y otras formas de coerción. Sus consecuencias pueden ser físicas, mentales, sexuales, reproductivas y económicas, afectando no solo a la víctima directa, sino también a hijos, familias y comunidades (OMS, 2024).

Para el presente trabajo es importante distinguir entre causas estructurales y factores situacionales o desencadenantes de corto plazo. Las causas estructurales hacen referencia a factores de fondo que aumentan el riesgo de violencia, como la desigualdad de género, las normas sociales que toleran la violencia, el control dentro de la pareja o la dependencia económica. Esta distinción es coherente con los enfoques de salud pública, que entienden la violencia de pareja como un fenómeno multicausal, asociado a factores individuales, relacionales, comunitarios y sociales, sin reducirlo a una única causa directa (CDC, 2024; OMS, 2019).

Los factores situacionales, en cambio, son acontecimientos o circunstancias que pueden aumentar temporalmente la tensión en un contexto donde ya existen factores de riesgo. No crean por sí mismos la violencia, pero pueden actuar como catalizadores o amplificadores de una situación previa. En este grupo pueden incluirse el estrés económico, los conflictos de pareja, el consumo problemático de alcohol o determinados eventos emocionalmente intensos (CDC, 2024). El interés de este TFG se sitúa en este segundo plano: analizar si determinados eventos futbolísticos pueden funcionar como shocks emocionales de corto plazo asociados a variaciones en las llamadas al 016.

Este enfoque exige una interpretación prudente. Si se observa un aumento de llamadas en días con ciertos eventos futbolísticos, no debe concluirse automáticamente que el fútbol sea la causa directa de la violencia. Más bien, el fútbol puede entenderse como un contexto social y emocional que, en determinadas circunstancias, coincide con una mayor demanda de atención o ayuda. Por ello, el trabajo habla de asociaciones y no de causalidad directa.

2.2 Eventos deportivos, frustración y comportamiento violento: evidencia internacional

La relación entre eventos deportivos y comportamiento violento ha sido analizada desde distintas disciplinas, como la economía, la criminología, la psicología social y la sociología del deporte. El interés común de estos estudios es entender cómo acontecimientos colectivos con fuerte carga emocional pueden afectar al comportamiento individual y social en el corto plazo.

Uno de los trabajos más relevantes es el de Card y Dahl (2011), que analiza la relación entre partidos de fútbol americano profesional y violencia doméstica en Estados Unidos. Los

autores estudian informes policiales de violencia en el hogar durante domingos de temporada regular y comparan los resultados deportivos con las expectativas previas del partido. Su principal hallazgo es que las derrotas inesperadas de equipos locales favoritos se asocian con un aumento de aproximadamente el 10% en la violencia en el hogar ejercida por hombres contra sus parejas. En cambio, las derrotas esperadas o las victorias inesperadas no presentan efectos comparables.

Este resultado es importante por dos motivos. En primer lugar, muestra que no basta con observar si hay partido o si el equipo pierde, sino que importa el componente emocional del resultado. La violencia aumenta especialmente cuando el resultado supone una pérdida respecto a las expectativas. En segundo lugar, el efecto se concentra en una ventana temporal cercana al final del partido, lo que refuerza la idea de que determinados eventos deportivos pueden actuar como shocks emocionales de corto plazo.

La lógica de Card y Dahl se apoya en ideas procedentes de la economía del comportamiento, especialmente en la teoría de la utilidad dependiente del punto de referencia. Según esta perspectiva, las personas no valoran los resultados únicamente en términos absolutos, sino en relación con lo que esperaban que ocurriera. Una derrota cuando se esperaba una victoria puede generar una pérdida psicológica mayor que una derrota esperada. Aplicado al fútbol, esto implica que no todas las derrotas tienen la misma carga emocional.

La literatura británica también ha analizado la relación entre fútbol y violencia doméstica, especialmente en el contexto de grandes torneos internacionales. Algunos estudios han observado aumentos de incidentes de violencia doméstica en días de partido de la selección inglesa, tanto tras victorias como tras derrotas, aunque con efectos más intensos en determinados resultados negativos. Estos trabajos subrayan que los partidos de alta audiencia pueden alterar rutinas domésticas, consumo de alcohol, reuniones sociales y estados emocionales, creando un contexto propicio para el aumento de conflictos en hogares donde ya existe riesgo previo.

En conjunto, la evidencia internacional sugiere que los eventos deportivos pueden tener efectos sociales que van más allá del estadio. El deporte no se interpreta aquí como causa estructural de la violencia, sino como un evento emocionalmente saliente que puede intensificar tensiones existentes. Esta idea es la base conceptual del presente estudio.

2.3 Fútbol, alcohol y violencia asociada al deporte

Una parte importante de la literatura sobre deporte y violencia destaca el papel del alcohol. El consumo de alcohol aparece de forma recurrente como factor de riesgo en la violencia interpersonal y en la violencia de pareja. La OMS lo incluye entre los factores asociados a la violencia contra las mujeres, y el CDC lo recoge también como factor de riesgo individual y comunitario en la violencia de pareja (CDC, 2024; OMS, 2024).

En el contexto deportivo, el alcohol puede estar vinculado a celebraciones, reuniones sociales, bares, desplazamientos, peñas o consumo doméstico durante los partidos. En muchos países, el fútbol se asocia a rutinas de ocio en las que el consumo de alcohol forma parte del ambiente del evento. Esto no implica que todos los aficionados consuman alcohol ni que el alcohol explique por sí solo la violencia, pero sí permite entender por qué determinados

eventos futbolísticos pueden coincidir con contextos de mayor riesgo.

Algunos estudios sobre violencia entre aficionados han señalado que los partidos de mayor rivalidad, las fases eliminatorias, las finales, los derbis y los encuentros con marcadores ajustados pueden generar mayor tensión y probabilidad de incidentes. Estos factores son relevantes para el presente trabajo porque conectan con varias de las variables construidas en la base de datos: derbis, eliminaciones, partidos de tensión alta, derrotas en descuento, remontadas sufridas y goles en contra en momentos finales.

El fútbol, además, tiene un componente identitario muy fuerte. Los equipos representan ciudades, provincias, barrios o comunidades. Esta identificación puede intensificar la reacción emocional ante determinados resultados. Una derrota en un partido ordinario puede no tener el mismo significado que una derrota en un derbi, una eliminación copera o un partido de final de temporada. Por ello, el presente trabajo no se limita a diferenciar entre victoria y derrota, sino que incorpora variables que intentan captar la intensidad emocional y competitiva del evento deportivo.

Es importante señalar que este TFG no dispone de datos directos sobre consumo de alcohol. Por tanto, el alcohol no se incorpora como variable empírica en los modelos. Sin embargo, se considera en el marco teórico como un posible mecanismo intermedio que puede ayudar a interpretar la relación entre fútbol, ocio, emociones y llamadas al 016. Esta limitación se tendrá en cuenta al discutir los resultados.

2.4 Evidencia empírica previa sobre resultados deportivos y violencia doméstica

La evidencia empírica más cercana a este trabajo se puede agrupar en tres líneas. La primera estudia el efecto de resultados deportivos sobre violencia doméstica. La segunda analiza el papel de grandes eventos deportivos en incidentes de violencia o criminalidad. La tercera se centra en los factores que aumentan la probabilidad de violencia en contextos de ocio, rivalidad o consumo de alcohol.

Dentro de la primera línea, Card y Dahl (2011) constituyen la referencia principal. Su estudio destaca porque utiliza datos policiales, resultados deportivos y expectativas de apuestas para identificar derrotas inesperadas. El elemento central no es simplemente perder, sino perder cuando el equipo era favorito. Esto permite interpretar el resultado como una pérdida emocional respecto a un punto de referencia previo. El hallazgo de un incremento de la violencia doméstica tras derrotas inesperadas ha convertido este trabajo en una referencia básica para analizar la relación entre deporte, emociones y violencia.

En Reino Unido, Kirby, Francis y O'Flaherty (2014) analizan la relación entre partidos de la selección inglesa en la Copa del Mundo y violencia doméstica. Sus resultados apuntan a incrementos de incidentes en días de partido, con diferencias según el resultado y el contexto del encuentro. Aunque el diseño y la fuente de datos son distintos a los de Card y Dahl, ambos enfoques coinciden en una idea: los eventos deportivos de gran relevancia social pueden coincidir con aumentos de violencia doméstica o demanda de servicios relacionados.

La segunda línea de literatura se centra en eventos deportivos como momentos de concentración social. Grandes partidos, derbis o fases finales pueden alterar patrones de

movilidad, ocio, consumo, convivencia y exposición mediática. Esta literatura no siempre se refiere específicamente a violencia doméstica, pero ayuda a entender por qué el deporte puede tener efectos sociales de corto plazo. La intensidad del evento, la rivalidad, la incertidumbre del resultado y el horario son elementos que pueden modificar el contexto en el que se producen comportamientos violentos o conflictos.

La tercera línea se refiere a factores asociados a violencia y agresividad en contextos deportivos. Algunos estudios sobre violencia entre aficionados destacan la importancia de la rivalidad, la competitividad del partido, la fase del torneo y la cercanía del marcador. Estos elementos son relevantes para este TFG porque permiten justificar variables como derbi, partido de tensión alta, eliminación, diferencia de goles, remontada sufrida o derrota en descuento.

A pesar de estas aportaciones, la evidencia para España es limitada. Existen datos oficiales sobre violencia de género y sobre llamadas al 016, pero no es habitual que se crucen con eventos deportivos a nivel territorial y temporal. Por ello, este trabajo aporta una aproximación empírica novedosa al caso español, utilizando datos de llamadas al 016 y una base propia de eventos futbolísticos desagregada territorial y temporalmente.

2.5. Aportación esperada del presente estudio al caso español

El presente trabajo contribuye a la literatura existente en varios aspectos:

Traslada al contexto español una pregunta que ha sido estudiada principalmente en Estados Unidos y Reino Unido: la posible relación entre eventos deportivos y violencia doméstica o demanda de ayuda asociada a violencia de género.

Utiliza como variable dependiente las llamadas pertinentes al Servicio 016. Esta fuente no mide directamente episodios de violencia, pero sí recoge demanda de información y apoyo institucional. Esto permite analizar una dimensión distinta a la denuncia policial o al registro judicial. En muchos casos, la llamada al 016 puede ser un paso previo, una consulta o una vía de orientación para víctimas o personas cercanas. Además, el patrón semanal del 016 muestra que las llamadas están condicionadas por dinámicas temporales propias del servicio, con mayor volumen al inicio de la semana y menor durante el fin de semana, lo que refuerza la necesidad de tratar las llamadas como indicador de demanda de atención y no como medición directa de la incidencia real de violencia (Observatorio Estatal de Violencia sobre la Mujer, 2022).

Incorpora una dimensión territorial provincial: Esta decisión es relevante porque los equipos de fútbol tienen una implantación local muy marcada. El impacto emocional de un resultado deportivo puede estar más concentrado en la provincia o ciudad del equipo que en el conjunto de la comunidad autónoma. Aunque la provincia sigue siendo un nivel agregado y no permite observar diferencias municipales, supone una aproximación territorial más precisa que el análisis autonómico.

No se limita a distinguir entre días con fútbol y días sin fútbol. La base futbolística permite construir variables más detalladas que captan distintos tipos de intensidad deportiva: derrotas, goles en contra, remontadas sufridas, derrotas en descuento, derbis, eliminaciones,

partidos nocturnos, rachas negativas y partidos de tensión alta. Esto permite comprobar si la posible asociación con las llamadas se debe a la actividad futbolística en general o a eventos más concretos.

Por último, el trabajo combina análisis descriptivo en Power BI con análisis estadístico en RStudio. Esta combinación permite, por un lado, visualizar patrones temporales, territoriales y deportivos de forma intuitiva y, por otro, contrastar posteriormente las relaciones observadas mediante modelos con controles. La finalidad no es establecer una prueba causal definitiva, sino aportar evidencia exploratoria y rigurosa sobre una cuestión socialmente relevante.

En síntesis, el marco teórico sugiere que los eventos deportivos pueden actuar como desencadenantes emocionales de corto plazo en determinados contextos. Sin embargo, la literatura también muestra que los efectos no son automáticos ni homogéneos: dependen del resultado, de las expectativas, de la importancia del partido, del contexto social y de factores no observados como el consumo de alcohol o la convivencia familiar. Por ello, el análisis empírico de este TFG se plantea desde una perspectiva prudente, centrada en asociaciones estadísticas entre actividad futbolística registrada y llamadas al 016.

3. Datos y fuentes de información

3.1 Servicio 016: bases de llamadas diarias y por franja horaria

El estudio se apoya en los datos de llamadas pertinentes al Servicio 016, teléfono de marcación abreviada que proporciona información, asesoramiento jurídico y atención psicosocial a víctimas de violencia de género y a su entorno. La utilización de esta fuente no implica observar directamente todos los episodios de violencia de género, sino aproximarse a la demanda de ayuda, información y asesoramiento registrada a través de este canal oficial. Por tanto, las llamadas al 016 se interpretan en este trabajo como un indicador indirecto de tensión o necesidad de asistencia relacionada con violencia de género.

La Delegación del Gobierno contra la Violencia de Género publica estadísticas del Servicio 016 con cierto nivel de agregación. Sin embargo, el objetivo de este trabajo exigía disponer de información más desagregada temporal y territorialmente, ya que se pretende analizar si determinados eventos deportivos pueden asociarse a variaciones de corto plazo en el volumen de llamadas. Por este motivo, se realizaron varias gestiones con el Ministerio de Igualdad, el Instituto de las Mujeres y distintas comunidades autónomas para solicitar datos con mayor nivel de detalle. Tras este proceso, se obtuvieron dos ficheros Excel con información sobre llamadas pertinentes al 016 desde 2021, ambos desagregados por provincia, aunque con distinta estructura temporal.

La primera fuente corresponde a una base diaria por provincia. Este fichero recoge, para cada provincia española y día del periodo analizado, el número de llamadas pertinentes registradas. Las variables originales principales son Año, Mes, Día, Provincia y Llamadas pertinentes. Esta base constituye la fuente principal del análisis diario, ya que permite observar la evolución de las llamadas en cada provincia y relacionarla posteriormente con la actividad futbolística registrada en la misma fecha.

La segunda fuente corresponde a una base por franja horaria. Este fichero recoge, para cada provincia y mes, el número de llamadas pertinentes distribuidas por tramos horarios de una hora. Las variables principales son Año, Mes, Provincia, Franja horaria y Llamadas pertinentes. A diferencia de la base diaria, esta fuente no permite observar el número de llamadas en una hora concreta de un día concreto, sino el volumen mensual acumulado por franja horaria. Por este motivo, su uso en el trabajo tiene un carácter complementario y exploratorio, orientado a estudiar patrones horarios agregados.

Aunque los datos originales del 016 cubren el conjunto de provincias españolas, el análisis empírico del presente trabajo se centra en Andalucía y la Región de Murcia, por ser las comunidades autónomas asignadas al ámbito de estudio. Para ello, se incorporó una clasificación territorial que permite identificar la comunidad autónoma y provincia correspondiente a cada observación.

La importancia del Servicio 016 como fuente de información se refleja en su volumen de actividad. Según los datos oficiales consultados, en 2025 se atendieron 110.595 consultas, con una media de 303 llamadas diarias. Además, la mayoría de las llamadas fueron realizadas por las propias víctimas, lo que refuerza la utilidad de esta fuente como aproximación a situaciones de demanda de ayuda. En todo caso, el análisis debe interpretarse con prudencia: un aumento de llamadas no equivale necesariamente a un aumento exacto de episodios de violencia, sino a un incremento en las consultas o solicitudes de apoyo registradas a través del servicio.

3.2 Base futbolística: registro de partidos y goles

Para medir la actividad futbolística y aproximar la intensidad emocional asociada a los partidos, se construyó una base de datos propia a partir de un registro detallado de goles de equipos de Primera División, Segunda División, Copa del Rey y competiciones europeas, correspondiente a las temporadas 2019-20 a 2024-25. El ámbito territorial se limitó a los equipos vinculados a Andalucía y la Región de Murcia, de acuerdo con el enfoque territorial del trabajo.

La base futbolística procede de un proceso de recopilación desarrollado mediante scripts en Python. Dado que no se disponía de una base ya preparada con el nivel de detalle requerido para el trabajo, se utilizó una API especializada en datos futbolísticos para descargar la información equipo por equipo y temporada por temporada. Es importante señalar que el script utilizado no construía toda la base completa en una única ejecución. Cada ejecución descargaba los datos correspondientes a un equipo concreto y a una temporada determinada. Por ello, antes de ejecutar el script se modificaban manualmente algunos parámetros, entre ellos el nombre del equipo de referencia, el identificador oficial del equipo en la API, la temporada analizada y otras variables auxiliares, como la división o la comunidad autónoma.

El proceso de recopilación combinaba información del calendario de partidos y de los eventos de gol. Para cada equipo y temporada se obtenían datos como la fecha del partido, la jornada, la competición, el equipo local, el equipo visitante, la hora oficial de inicio y el resultado final. Posteriormente, para cada partido se recogían los eventos de gol, incluyendo el minuto, el equipo anotador y la identificación del gol como favorable o contrario al equipo de referencia.

Cada ejecución del script generaba un fichero Excel individual, que posteriormente fue integrado manualmente en un Excel maestro utilizado como base futbolística final del trabajo.

La unidad original de observación de esta base no es el partido, sino el evento de gol. Por tanto, un mismo partido aparece repetido tantas veces como goles se marcaron en él. Por ejemplo, un partido con resultado 2-3 genera cinco filas, una por cada gol. Cada fila contiene información del partido completo, como fecha, temporada, equipo local, equipo visitante, resultado, competición y hora de inicio, junto con información específica del gol, como minuto, franja horaria del gol, equipo marcador y clasificación como gol a favor o en contra del equipo de referencia.

Esta estructura resulta útil para analizar la dimensión temporal de los goles, pero presenta una limitación importante: los partidos finalizados 0-0 no quedan representados en el registro. Por este motivo, la base no debe interpretarse como un calendario completo de todos los partidos disputados, sino como un registro de partidos con eventos de gol. Esta aclaración es relevante para la interpretación del análisis, ya que algunas temporadas pueden mostrar menos partidos observados que los realmente disputados.

Tras la consolidación del Excel maestro, se realizó una revisión de calidad de la base. Esta revisión consistió en comprobar, por equipo, temporada y competición, si existían registros coherentes con la participación real de cada equipo. En esta fase se detectó una ausencia relevante: la temporada 2020-21 de la UD Almería en Segunda División no estaba incluida en la base, aunque el equipo sí disputó dicha competición. Esta ausencia no podía explicarse por partidos finalizados 0-0, ya que no faltaban algunos registros aislados, sino prácticamente toda la temporada liguera. Para corregirlo, se realizó una búsqueda específica mediante Deep Research, contrastando los partidos y goles de la UD Almería en Segunda División 2020-21 y reconstruyendo los registros necesarios con la misma estructura del Excel maestro. Una vez completada esta información, se incorporó a la base general para mantener la coherencia del panel de equipos y temporadas.

A partir del registro original de goles se generó posteriormente una tabla a nivel partido, necesaria para evitar que un mismo encuentro apareciera repetido varias veces en los análisis diarios. Esta transformación y la construcción de variables deportivas específicas se explican con mayor detalle en el apartado metodológico.

3.3 Tablas auxiliares y preparación de datos para Power BI y RStudio

Para integrar las fuentes de llamadas y fútbol se construyó un modelo de datos en Power BI compuesto por tablas de hechos y tablas auxiliares. Las tablas de hechos principales son la base diaria de llamadas, la base de llamadas por franja horaria, el registro de goles y la tabla de partidos derivada del registro futbolístico.

También se incorporaron tablas auxiliares de calendario, territorio y franjas horarias. La tabla calendario permite ordenar temporalmente la información y relacionar las llamadas diarias con los eventos futbolísticos por fecha. Las tablas territoriales permiten asociar cada equipo de referencia a su provincia y comunidad autónoma, lo que hace posible cruzar los resultados deportivos con las llamadas al 016 a nivel provincial. Por su parte, el maestro de franjas horarias permite armonizar las franjas de goles, partidos y llamadas en el análisis horario

agregado.

A partir de este modelo se prepararon dos bases finales para el análisis estadístico en RStudio. La primera tiene una estructura diaria, con una fila por provincia y fecha, y permite estudiar la relación entre llamadas al 016 y actividad futbolística registrada en el mismo día. La segunda tiene una estructura mensual por provincia y franja horaria, y permite explorar patrones horarios agregados. La construcción concreta de estas bases, así como las variables deportivas incluidas en cada una, se desarrolla en el apartado de metodología.

En conjunto, el proceso de preparación de datos permitió transformar tres fuentes iniciales (llamadas diarias, llamadas por franja horaria y registro de goles) en un modelo analítico integrado. El resultado son dos niveles de análisis complementarios: un análisis diario, más adecuado para estudiar variaciones de corto plazo por fecha y provincia, y un análisis horario agregado, útil para explorar patrones temporales dentro del día.

3.4 Limitaciones y nivel de desagregación

Las bases utilizadas presentan varias limitaciones que deben tenerse en cuenta en la interpretación de los resultados. Aunque el objetivo inicial era disponer de llamadas al 016 por provincia, día y franja horaria, la información finalmente obtenida se divide en dos fuentes: una base diaria por provincia y una base por franja horaria agregada a nivel mensual. Por ello, el análisis diario permite estudiar asociaciones entre actividad futbolística registrada y llamadas en una misma fecha y provincia, mientras que el análisis horario debe interpretarse únicamente como una aproximación exploratoria basada en patrones mensuales por franja.

Además, la base futbolística se construye a partir de eventos de gol, por lo que los partidos finalizados 0-0 no aparecen en el registro. Esto implica que las variables deportivas no representan todos los partidos disputados, sino aquellos encuentros con goles registrados. Por esta razón, a lo largo del trabajo se emplean expresiones como “actividad futbolística registrada” o “partidos con eventos de gol”, evitando interpretar la base como un calendario completo de partidos.

Por otro lado, el cruce territorial se realiza a nivel provincial. Este nivel de desagregación permite un análisis más preciso que el autonómico, pero no capta diferencias dentro de una misma provincia. El impacto emocional de un resultado deportivo puede concentrarse especialmente en la ciudad del equipo o en su área de influencia más directa, algo que no puede observarse con los datos disponibles.

Finalmente, el diseño del estudio permite analizar asociaciones estadísticas, pero no establecer relaciones causales directas. Las variaciones en las llamadas al 016 pueden estar influidas por otros factores no observados, como consumo de alcohol, clima, eventos sociales, noticias recientes, festivos o cambios en la visibilidad pública de la violencia de género. En consecuencia, los resultados deben interpretarse como evidencia exploratoria sobre la posible relación entre determinados contextos futbolísticos y la demanda de atención registrada en el 016, no como una prueba causal definitiva.

A pesar de estas limitaciones, la combinación de llamadas diarias por provincia, llamadas

mensuales por franja horaria y una base propia de eventos futbolísticos permite construir un análisis aplicado al contexto español con un nivel de detalle suficiente para complementar la evidencia internacional existente.

4. Metodología

4.1 Diseño general

El objetivo metodológico del trabajo es analizar si determinados eventos futbolísticos pueden estar asociados a variaciones de corto plazo en el volumen de llamadas pertinentes al Servicio 016. Para ello, se adopta un enfoque cuantitativo basado en tres elementos principales: la integración de bases de datos procedentes de fuentes distintas, el análisis descriptivo mediante Power BI y la estimación posterior de modelos estadísticos en RStudio.

El trabajo no pretende demostrar una relación causal directa entre fútbol y violencia de género, ya que los datos disponibles no permiten observar episodios individuales ni vincular una llamada concreta con un partido específico. El planteamiento es más prudente: estudiar si existe una asociación estadística entre la actividad futbolística registrada y el número de llamadas al 016, prestando especial atención a contextos deportivos que podrían generar mayor carga emocional, como derrotas, goles tardíos en contra, remontadas sufridas, derbis, eliminaciones o partidos de alta tensión.

La metodología se organiza en cinco fases. Primero se integran las fuentes de llamadas y fútbol en un modelo común. Después, se construyen variables deportivas, temporales y territoriales adaptadas a la pregunta de investigación. Tras esto, se realiza un análisis descriptivo en Power BI para explorar patrones generales. Posteriormente se preparan dos bases específicas para RStudio: una diaria por provincia y otra mensual por provincia y franja horaria. Finalmente, se estiman modelos estadísticos para evaluar la asociación entre llamadas al 016 y variables futbolísticas, incorporando controles temporales y territoriales.

4.2 Unidad de análisis e integración de las bases

La integración de las fuentes se diseñó teniendo en cuenta el distinto nivel de desagregación disponible. La base diaria del 016 permite trabajar con una unidad de análisis provincia-fecha, mientras que la base por franjas horarias solo permite trabajar a nivel provincia-mes-franja. Por este motivo, el estudio se organiza en dos niveles complementarios: un análisis diario, que constituye la parte principal del trabajo, y un análisis horario agregado, de carácter exploratorio.

En el análisis diario, cada observación representa una provincia en una fecha concreta. Para cada provincia y día se dispone del número de llamadas pertinentes al 016 y de las variables futbolísticas correspondientes a los equipos asociados a esa provincia. Esta estructura permite estudiar si los días con actividad futbolística registrada, y especialmente los días con determinados resultados deportivos, presentan un volumen de llamadas diferente al resto.

En el análisis por franjas, cada observación representa una provincia, un mes y una franja horaria. Esta base permite comparar el volumen mensual de llamadas acumuladas en cada

tramo horario con la distribución temporal de partidos y goles. Debido a que no se dispone de llamadas por hora y día, este análisis no permite vincular una llamada a un partido concreto, pero sí permite estudiar si ciertos tramos horarios con mayor exposición futbolística presentan patrones diferenciados de llamadas.

La integración territorial se realiza a nivel provincial. Para ello, cada equipo de referencia se asocia a una provincia y comunidad autónoma. Esta decisión permite cruzar los eventos futbolísticos con las llamadas del 016 utilizando la provincia como unidad territorial común. Aunque este nivel no capta diferencias dentro de una misma provincia, es el nivel más desagregado disponible para las llamadas y permite un análisis más preciso que el autonómico.

4.3 Construcción de variables deportivas

A partir del registro de goles se construyó una tabla agregada a nivel partido. Esta transformación fue necesaria porque la base futbolística original está estructurada a nivel de gol, por lo que un mismo partido aparece repetido tantas veces como goles se marcaron. Para evitar duplicidades en el análisis diario, cada partido del equipo de referencia se resumió en una única fila.

La primera familia de variables recoge el resultado deportivo básico del encuentro. Para cada partido se calcularon los goles a favor del equipo analizado, los goles en contra, el total de goles y la diferencia de goles. A partir de estas variables se clasificó el resultado como victoria, empate o derrota. Esta clasificación permite analizar si los días con derrotas, victorias o empates presentan diferencias en el volumen de llamadas al 016.

La segunda familia de variables aproxima la intensidad emocional del resultado deportivo. Se construyeron variables para identificar derrotas tardías, derrotas en descuento, derrotas abultadas, remontadas sufridas, goles en contra tempranos, derrotas encajando primero, derrotas como local o visitante y derrotas con racha previa negativa. Estas variables buscan distinguir derrotas potencialmente más frustrantes o emocionalmente intensas de derrotas ordinarias.

La tercera familia de variables recoge el contexto competitivo del partido. Se identificaron derbis, eliminaciones, partidos en las últimas tres jornadas de liga, partidos nocturnos y partidos de alta intensidad. A partir de estos elementos se creó también una variable de partido de tensión alta, que agrupa encuentros con características especialmente relevantes desde el punto de vista competitivo o emocional. Posteriormente, se construyó una variable específica de derrota en partido de tensión alta, para aislar aquellos casos en los que un contexto relevante coincide además con un resultado negativo.

Estas variables no se interpretan como causas directas de violencia de género, sino como indicadores de exposición deportiva o emocional. Su objetivo es permitir contrastar empíricamente si determinados contextos futbolísticos coinciden con variaciones en la demanda de atención registrada a través del 016.

4.4 Bases finales de análisis

Para el análisis estadístico se prepararon dos bases finales: Dataset_R_Diario y

Dataset_R_Franjas.

La primera es Dataset_R_Diario, con una fila por provincia y fecha. Esta tabla contiene el número de llamadas pertinentes al 016 y las variables futbolísticas agregadas para esa misma provincia y día. En las fechas sin actividad futbolística registrada, las variables deportivas toman valor cero. Esta estructura evita duplicar el número de llamadas cuando en una misma provincia hay más de un partido en el mismo día y permite trabajar con una unidad de análisis clara: provincia-día.

La segunda es Dataset_R_Franjas, con una fila por provincia, mes y franja horaria. Esta tabla permite explotar la información horaria agregada del 016. En ella se distinguen tres tipos de exposición futbolística: las variables de goles en franja, que recogen los goles registrados en cada tramo horario; las variables durante el partido, que contabilizan los partidos cuya hora de inicio y fin se solapa con la franja analizada; y las variables de franja posterior, que asignan cada partido a la hora inmediatamente posterior a su finalización.

Esta distinción permite comparar tres posibles mecanismos temporales: que las llamadas se relacionen con el momento en que se están disputando los partidos, con los goles ocurridos en una franja concreta o con el periodo posterior al resultado final. No obstante, debido a la agregación mensual de la base horaria del 016, este segundo nivel de análisis se interpreta como exploratorio.

4.5 Análisis descriptivo en Power BI

El primer bloque analítico se desarrolla en Power BI. Su objetivo es explorar visualmente las principales relaciones entre llamadas al 016 y actividad futbolística antes de proceder a la estimación estadística. Este análisis descriptivo permite identificar patrones temporales, diferencias territoriales, valores atípicos y posibles relaciones que posteriormente pueden contrastarse en RStudio.

Primero se analiza la evolución temporal de las llamadas al 016. Para ello se construyen gráficos de líneas con el número de llamadas por día y provincia, diferenciando entre días con y sin actividad futbolística registrada. También se analiza la evolución por mes, día de la semana y tipo de día, distinguiendo entre semana y fin de semana. Esta separación es importante porque la distribución de llamadas y la programación de partidos no son homogéneas a lo largo de la semana.

Más adelante, se comparan las llamadas según el resultado deportivo. Se utilizan gráficos de barras, tablas y visualizaciones de distribución para comparar el volumen medio de llamadas en días con victorias, empates y derrotas. Además, se analizan variables más específicas, como derrotas tardías, derrotas en descuento, derrotas abultadas, remontadas sufridas, derrotas como local y derrotas en partidos de tensión alta. Esta parte del análisis permite observar si los días asociados a resultados deportivos negativos presentan niveles de llamadas superiores a otros días.

Luego se desarrolla un análisis territorial: El uso de la provincia como dimensión permite comparar patrones entre las provincias de Andalucía y la Región de Murcia. Este análisis es especialmente relevante porque el impacto emocional de un equipo puede estar concentrado

en su provincia de referencia. Los mapas, tablas y gráficos por provincia permiten observar si existen diferencias territoriales en la relación entre fútbol y llamadas al 016.

Después, se exploran las variables de intensidad deportiva. En Power BI se visualizan las llamadas medias asociadas a partidos de alta intensidad, derbis, eliminaciones, últimas jornadas, rachas de derrotas y partidos nocturnos. También se utiliza la variable de desencadenante principal como herramienta descriptiva para resumir el tipo de evento deportivo predominante en cada partido. Esta variable no se plantea como variable principal para los modelos estadísticos, ya que agrupa distintos fenómenos en una clasificación jerárquica, pero resulta útil para elaborar gráficos explicativos.

Por último, se realiza una exploración de las franjas horarias. Se comparan las llamadas mensuales por franja con goles ocurridos en la franja, partidos en juego durante la franja y partidos finalizados en la hora anterior. Esta distinción permite analizar si el patrón horario de llamadas se aproxima más al desarrollo del partido, al momento de los goles o al periodo posterior al encuentro.

4.6 Análisis estadístico y econométrico en RStudio

El segundo bloque analítico se realiza en RStudio. Su objetivo es evaluar de forma más sistemática la asociación entre las variables futbolísticas construidas y las llamadas al 016, incorporando controles temporales y territoriales. El análisis se desarrolla de forma separada para la base diaria y para la base por franjas, debido a que ambas tienen distinto nivel de desagregación.

El análisis diario utiliza como unidad de observación la provincia-fecha. La variable dependiente es el número de llamadas pertinentes al 016 registradas en una provincia y día determinados. Las principales variables explicativas son las variables futbolísticas agregadas para esa misma provincia y fecha: número de partidos registrados, derrotas, victorias, empates, goles a favor, goles en contra, diferencia de goles, derrotas tardías, derrotas en descuento, derrotas abultadas, remontadas sufridas, derbis, eliminaciones, partidos nocturnos, rachas previas de derrotas y partidos de tensión alta.

Como primer paso, se calculan estadísticos descriptivos y matrices de correlación. Estas matrices permiten identificar qué variables deportivas presentan mayor asociación lineal con el volumen de llamadas. No obstante, la interpretación de las correlaciones se realiza con cautela, ya que muchas variables deportivas están relacionadas entre sí. Por ejemplo, una derrota puede coincidir con goles en contra, diferencia de goles negativa, derrota local o derrota en partido de tensión alta. Por ello, las correlaciones se utilizan como herramienta exploratoria, no como prueba definitiva.

Posteriormente se realizan comparaciones de medias. Se comparan las llamadas medias en días con y sin actividad futbolística registrada, días con y sin derrotas, días con derrotas tardías frente a días sin ellas, y días con eventos de alta tensión frente al resto. Para estas comparaciones se utilizan pruebas t o análisis de varianza, según el número de grupos comparados. Además, se realizan comparaciones separadas para entre semana y fin de semana, dado que la distribución de partidos y llamadas puede diferir entre ambos tipos de día.

A continuación, se estiman modelos de regresión. La especificación básica toma la siguiente forma:

$$\text{Llamadas}_{it} = \beta_0 + \beta_1 \text{Derrotas}_{it} + \beta_2 \text{Victorias}_{it} + \beta_3 \text{GolesContra}_{it} + \beta_4 \text{FinSemana}_t + \delta_i + \gamma_t + \varepsilon_{it}$$

donde i representa la provincia y t representa el día. La variable dependiente es el número de llamadas al 016. Las variables deportivas recogen la actividad futbolística registrada en esa provincia y día. δ_i representa efectos fijos provinciales, que controlan diferencias estructurales entre territorios, y γ_t recoge controles temporales, como mes, año o día de la semana. Esta especificación permite comparar variaciones dentro de cada provincia, reduciendo el riesgo de que las diferencias territoriales constantes sesguen los resultados.

Además del modelo básico, se estiman modelos alternativos por bloques de variables. Un primer bloque analiza el resultado deportivo general, incluyendo victorias, derrotas y diferencia de goles. Un segundo bloque se centra en derrotas especialmente intensas, como derrotas tardías, derrotas en descuento, remontadas sufridas o derrotas abultadas. Un tercer bloque recoge el contexto competitivo, incluyendo derbis, eliminaciones, últimas jornadas y partidos de tensión alta. Esta estrategia evita introducir simultáneamente demasiadas variables altamente correlacionadas y facilita la interpretación de los coeficientes.

Dado que la variable dependiente es un recuento de llamadas, también se estiman modelos para datos de conteo, como Poisson y binomial negativa. Estos modelos resultan adecuados cuando la variable explicada toma valores enteros no negativos y puede presentar sobredispersión. La comparación entre regresión lineal, Poisson y binomial negativa permite evaluar la robustez de los resultados ante distintas especificaciones.

Asimismo, se exploran efectos temporales retardados. Para ello se incorporan variables futbolísticas rezagadas uno o varios días, con el objetivo de comprobar si los posibles efectos asociados al fútbol se concentran en el mismo día del partido o se prolongan a los días posteriores. Esta parte del análisis es relevante porque la literatura previa sugiere que algunos efectos pueden aparecer después del encuentro y no necesariamente durante el partido.

El análisis por franjas horarias se realiza con la base Dataset_R_Franjas. En este caso, la unidad de observación es provincia-mes-franja horaria y la variable dependiente es el número de llamadas acumuladas en esa franja durante el mes. Debido al nivel de agregación mensual, este análisis no se interpreta como causal ni como una medición exacta del efecto horario de partidos concretos. Su finalidad es exploratoria: comparar si las franjas con mayor presencia de partidos en juego, goles o periodos posteriores al partido coinciden con un mayor volumen de llamadas acumuladas.

Finalmente, como análisis complementario, se plantea un ejercicio predictivo exploratorio. Para ello se puede construir una variable binaria que identifique los días con un volumen de llamadas superior a un determinado percentil, por ejemplo el percentil 75. A partir de esta variable se pueden estimar modelos de clasificación, como regresión logística o bosques aleatorios, incorporando variables temporales, territoriales y deportivas. El objetivo de este ejercicio no es sustituir el análisis econométrico principal, sino comprobar si las variables futbolísticas aportan información adicional para identificar días de mayor presión sobre el

Servicio 016.

4.7 Fiabilidad e interpretación de resultados

La interpretación de los resultados se realiza con cautela por tres motivos principales:

- La base futbolística no recoge partidos finalizados 0-0, por lo que las variables deportivas deben entenderse como actividad futbolística registrada y no como calendario completo de partidos.
- El análisis por franjas tiene carácter exploratorio, ya que las llamadas por franja están agregadas a nivel mensual.
- Pueden existir factores no observados que influyan en el volumen de llamadas, como consumo de alcohol, clima, festivos, eventos sociales, noticias concurrentes o cambios en la visibilidad pública de la violencia de género.

Para reforzar la fiabilidad del análisis, se plantean varias comprobaciones:

- Se comparan resultados con y sin controles temporales y territoriales.
- Se estiman modelos separados para entre semana y fin de semana.
- Se analizan distintas familias de variables deportivas por separado, para evitar problemas derivados de la alta correlación entre ellas.
- Se contrastan los resultados mediante distintas especificaciones estadísticas, incluyendo regresión lineal y modelos de conteo.

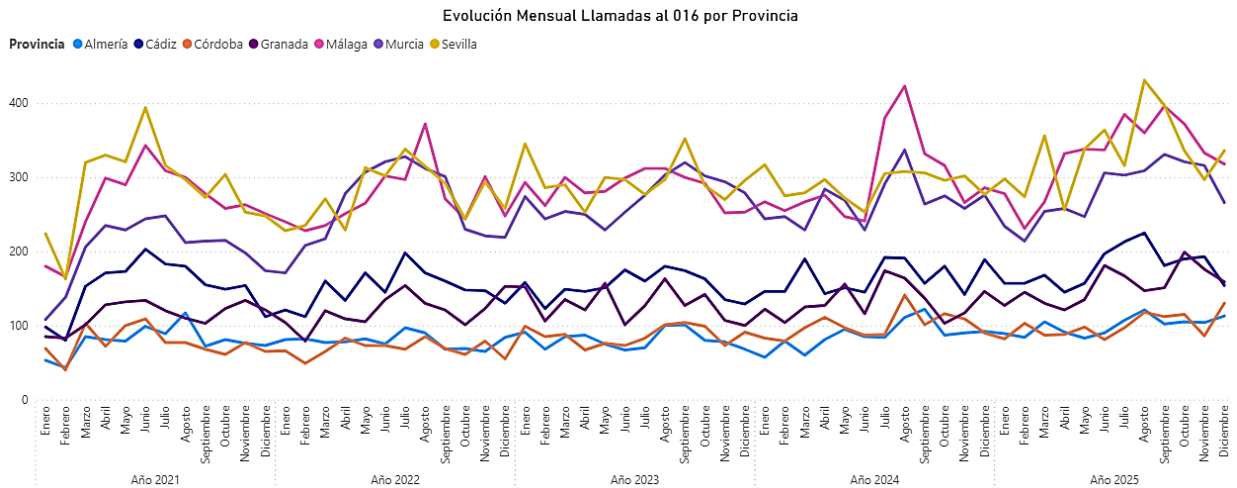
En conjunto, la metodología combina la construcción de una base de datos propia, el análisis visual en Power BI y la estimación estadística en RStudio. Esta combinación permite abordar la pregunta de investigación desde distintos niveles de análisis, aprovechando la información diaria, territorial y horaria disponible, pero manteniendo una interpretación prudente acorde con las limitaciones de los datos.

5. Análisis Descriptivo - Power BI

5.1 Análisis General

5.1.1 Evolución general de las llamadas al 016

Antes de analizar la posible relación entre actividad futbolística y llamadas al 016, resulta necesario describir el comportamiento general de la variable dependiente. En primer lugar, se representa la evolución mensual de las llamadas pertinentes al 016 por provincia para el periodo disponible (2021-2025). Esta visualización permite observar diferencias territoriales claras y una evolución temporal relativamente estable, aunque con variaciones mensuales.

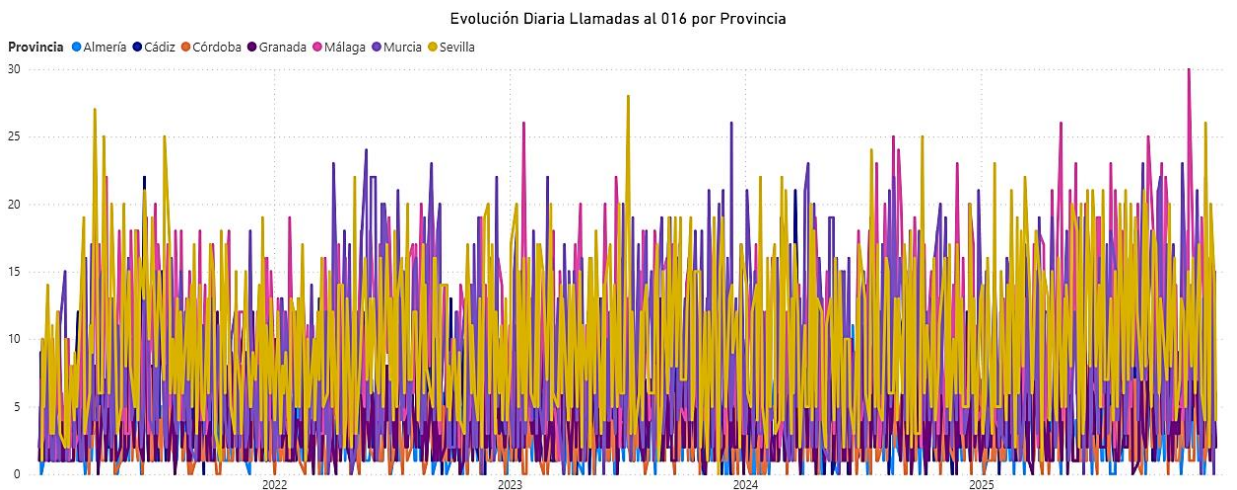


El gráfico mensual muestra que las provincias con mayor volumen de llamadas son Sevilla, Málaga y Murcia. Córdoba y Almería son las provincias con menor número de llamadas del periodo seleccionado.

También se aprecia que las llamadas no siguen una tendencia completamente uniforme. Existen meses con incrementos puntuales, especialmente en Sevilla, Málaga y Murcia, aunque no se observa una ruptura estructural clara en el conjunto de la serie. En términos generales, el volumen de llamadas parece oscilar alrededor de niveles relativamente estables, con aumentos puntuales en algunos meses de 2024 y 2025.

En un segundo nivel aparecen Cádiz y Granada, con cifras mensuales intermedias. Córdoba y Almería presentan los niveles más reducidos dentro del conjunto analizado. Esta distribución territorial es relevante para el análisis posterior, ya que confirma que las provincias no parten del mismo nivel de llamadas y que será necesario tener en cuenta estas diferencias en la interpretación de los resultados.

Para complementar esta visión agregada, se representa también la evolución diaria de las llamadas al 016 por provincia. Esta visualización muestra una mayor volatilidad, como es esperable al trabajar con datos diarios.



El gráfico diario permite observar que las llamadas presentan fuertes variaciones de un día a otro. Esta volatilidad es visible en todas las provincias, aunque con distinta intensidad según el volumen medio de cada territorio. Sevilla, Málaga y Murcia muestran picos diarios más

elevados, mientras que Almería, Córdoba y Granada se concentran generalmente en valores más bajos.

La elevada variabilidad diaria justifica la necesidad de complementar el análisis visual con medidas agregadas y modelos estadísticos. A simple vista, resulta difícil identificar patrones claros en la serie diaria, especialmente cuando se representan simultáneamente varias provincias. Por este motivo, en los siguientes apartados se utilizan comparaciones por tipo de día, provincia y presencia de actividad futbolística registrada.

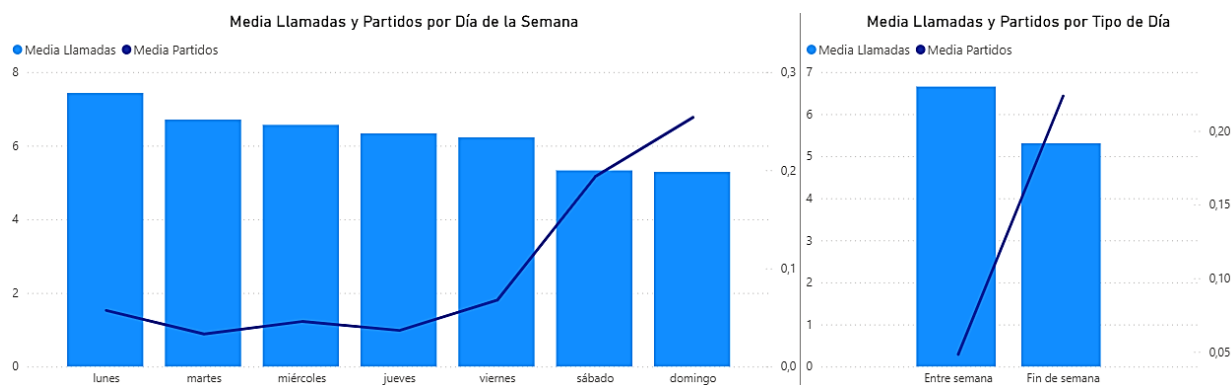
5.1.2 Diferencias por día de la semana y tipo de día

Uno de los aspectos más relevantes para el análisis es la diferencia entre días laborables y fines de semana. Esta distinción es especialmente importante porque la programación futbolística se concentra en gran medida durante el fin de semana. Por el contrario, las llamadas al 016 presentan una dinámica distinta.

El gráfico de llamadas y partidos por día de la semana muestra dos patrones claramente diferenciados. Por un lado, las llamadas medias al 016 son más elevadas al comienzo de la semana, especialmente los lunes, con una media de 7,4 llamadas. A partir de ahí, la media desciende progresivamente durante la semana, situándose en 5,3 llamadas durante sábado y domingo.

Por otro lado, la actividad futbolística registrada se concentra principalmente en sábado y domingo, coincidiendo con la estructura habitual del calendario deportivo. Mientras que entre semana la media de partidos es reducida, el sábado y el domingo muestran los valores más elevados.

Esta diferencia es fundamental desde el punto de vista metodológico. Si se comparasen directamente los días con fútbol frente a los días sin fútbol sin tener en cuenta el tipo de día, se correría el riesgo de confundir el efecto del calendario semanal con el posible efecto de la actividad futbolística. Es decir, los días con más partidos coinciden con días en los que, de forma general, las llamadas medias tienden a ser menores. Por ello, el análisis posterior diferencia sistemáticamente entre “Entre semana” y “Fin de semana”.



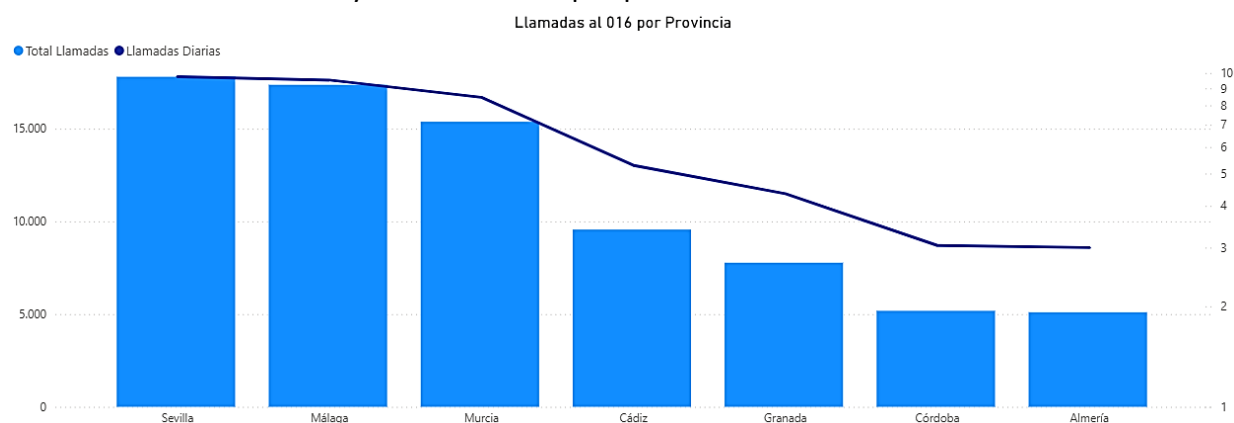
La comparación agregada por tipo de día confirma esta idea. Las llamadas medias son superiores entre semana, con una media de 6,7 llamadas, frente a 5,3 llamadas en fin de semana. En cambio, la actividad futbolística media es claramente mayor durante el fin de semana.

Este resultado refuerza la necesidad de interpretar con cautela cualquier comparación simple entre días con y sin fútbol. No basta con observar si hay más o menos llamadas en días con actividad futbolística registrada; es necesario analizar si esa diferencia se mantiene dentro de cada tipo de día.

En consecuencia, las hipótesis descriptivas posteriores se analizarán, siempre que sea posible, diferenciando entre días laborables y fines de semana. Esta separación permite aproximarse mejor a la relación entre fútbol y llamadas al 016, reduciendo el riesgo de atribuir al fútbol diferencias que pueden explicarse por la propia estructura semanal de las llamadas.

5.1.3 Distribución territorial de las llamadas

Además de las diferencias temporales, el análisis descriptivo muestra importantes diferencias territoriales en el volumen de llamadas. Para observarlo de forma más clara, se representa el número total de llamadas y la media diaria por provincia.



El gráfico por provincia confirma que Sevilla, Málaga y Murcia son las provincias con mayor volumen de llamadas dentro del ámbito analizado. Sevilla presenta el mayor número total, aunque Málaga se sitúa en un nivel muy cercano. Murcia aparece como la tercera provincia con mayor volumen, seguida por Cádiz y Granada. Córdoba y Almería presentan los valores más bajos del conjunto.

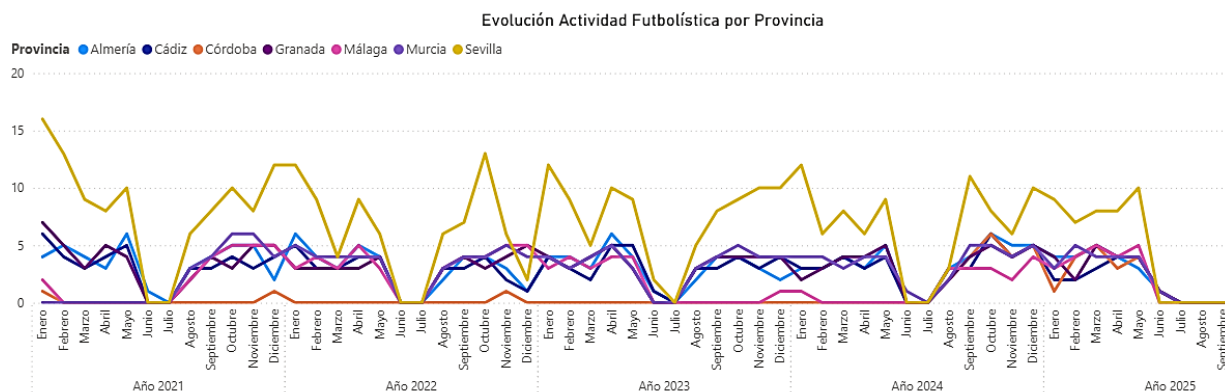
En términos de media diaria, el patrón es similar. Sevilla y Málaga se sitúan en 9,8 y 9,5 llamadas diarias respectivamente, Murcia 8,5, Cádiz en 5,3, Granada en 4,4, y Córdoba y Almería alrededor de 3 llamadas diarias. Estas diferencias pueden responder a múltiples factores, como el tamaño poblacional, la estructura territorial, la visibilidad del servicio, las dinámicas sociales o la propia distribución de la demanda de atención.

Desde el punto de vista del estudio, esta heterogeneidad territorial implica que no todas las provincias son directamente comparables en términos absolutos. Por ello, en el análisis estadístico posterior será importante controlar por provincia o comparar variaciones dentro de cada territorio.

También es relevante tener en cuenta que las provincias con mayor volumen de llamadas no coinciden necesariamente con las provincias con mayor intensidad futbolística en todos los momentos del periodo. Esta posible diferencia entre volumen de llamadas y exposición futbolística justifica el cruce posterior entre ambas dimensiones.

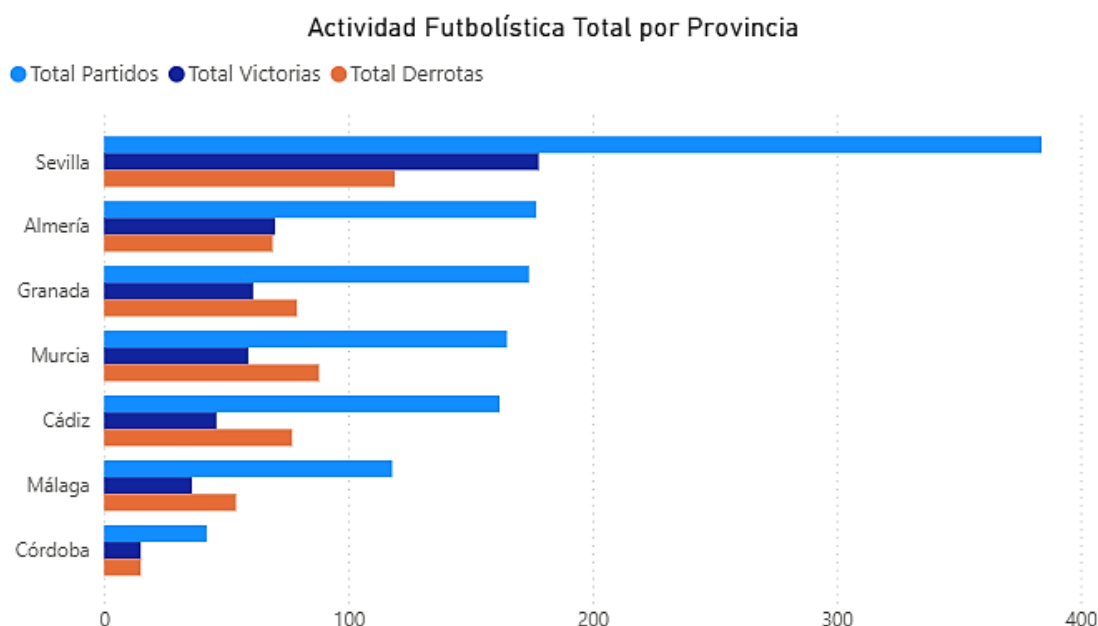
5.1.4 Actividad futbolística registrada por provincia

Una vez descrita la evolución de las llamadas, se analiza la distribución de la actividad futbolística registrada. Dado que la base futbolística procede de eventos de gol, debe recordarse que las cifras representan partidos con goles registrados y no el calendario completo de partidos disputados. Aun así, la variable permite aproximar la exposición futbolística de cada provincia durante el periodo de estudio.



La evolución mensual de la actividad futbolística muestra un patrón estacional claro. La actividad se concentra durante los meses habituales de competición y disminuye de forma notable en los meses de verano, especialmente en julio. Este comportamiento es coherente con el calendario futbolístico, en el que las temporadas se desarrollan principalmente entre agosto o septiembre y mayo o junio.

También se observan diferencias entre provincias. Sevilla presenta de forma recurrente una mayor actividad futbolística registrada, lo que resulta coherente con la presencia de dos equipos con participación continuada en competiciones nacionales y europeas durante buena parte del periodo. Otras provincias, como Almería, Granada, Cádiz, Málaga o Murcia, presentan niveles de actividad más moderados y sujetos a la participación de sus equipos en Primera, Segunda, Copa del Rey o playoff de ascenso. Córdoba muestra una actividad más reducida, en parte porque su presencia en las competiciones consideradas se concentra en menos temporadas.



El gráfico de actividad futbolística total por provincia confirma esta distribución. Sevilla es la

provincia con mayor número de partidos registrados, victorias y derrotas. Este resultado era esperable por la presencia simultánea del Sevilla FC y el Real Betis, ambos con continuidad en Primera División durante todo el periodo analizado y con participación adicional en competiciones como Copa del Rey o competiciones europeas.

Almería, Granada, Murcia y Cádiz presentan también un volumen relevante de actividad futbolística, aunque inferior al de Sevilla. Málaga y Córdoba muestran menor actividad total, lo que se relaciona con las temporadas en las que sus equipos no participaron en las competiciones incluidas en la base. En conjunto, estos resultados muestran que la exposición futbolística no está distribuida de forma homogénea entre provincias, lo que deberá tenerse en cuenta en el análisis de hipótesis.

La comparación entre llamadas y actividad futbolística permite identificar una primera cuestión metodológica relevante. Sevilla combina un volumen elevado de llamadas y una elevada exposición futbolística, pero esta coincidencia no implica por sí misma una relación entre fútbol y llamadas. Puede estar influida por otros factores estructurales, como población, peso urbano o características territoriales. Por ello, el análisis posterior no se limitará a comparar niveles absolutos entre provincias, sino que estudiará diferencias dentro de cada provincia y según el tipo de evento deportivo.

5.1.5 Síntesis del análisis descriptivo general

El análisis descriptivo inicial permite extraer varias conclusiones relevantes para orientar el resto del trabajo. En primer lugar, las llamadas al 016 presentan diferencias territoriales importantes. Sevilla, Málaga y Murcia concentran los mayores volúmenes, mientras que Córdoba y Almería presentan los niveles más bajos.

En segundo lugar, existe una fuerte diferencia temporal entre días laborables y fines de semana: las llamadas medias son más elevadas entre semana, mientras que la actividad futbolística se concentra en sábado y domingo. Esta diferencia obliga a separar ambos tipos de día en los análisis posteriores.

En tercer lugar, la actividad futbolística registrada muestra una distribución territorial desigual y un patrón estacional coherente con el calendario deportivo. Sevilla concentra la mayor exposición futbolística, seguida por provincias con equipos presentes en Primera, Segunda, Copa del Rey o playoff de ascenso durante distintas temporadas.

Por último, los gráficos generales muestran que no es suficiente analizar únicamente si hay fútbol o no, sino que resulta necesario distinguir entre tipos de resultado y contextos deportivos concretos. A partir de estas observaciones, el análisis descriptivo continúa con el contraste de hipótesis específicas basadas en las variables deportivas construidas: actividad futbolística, victorias, derrotas, número de derrotas, partidos de tensión alta, derrotas tardías, remontadas sufridas, derbis, eliminaciones, rachas negativas y goles en contra. Estas hipótesis se analizarán diferenciando entre “entre semana” y “fin de semana”, para evitar que los patrones propios del calendario semanal distorsionen la interpretación.

5.2 Análisis Diario - Contraste descriptivo de hipótesis

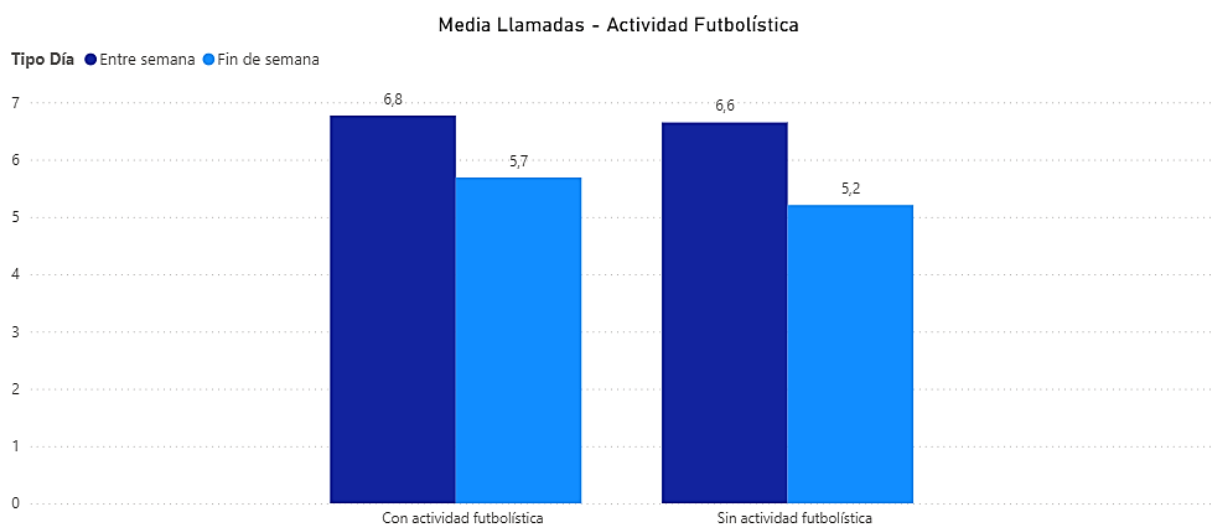
Una vez descritos los patrones generales de llamadas y actividad futbolística, el análisis se

centra en contrastar de forma descriptiva las principales hipótesis del trabajo. Para ello, se comparan las llamadas medias al 016 en función de distintas variables deportivas construidas a partir de la base de partidos. En todos los casos, el análisis se diferencia entre días de entre semana y fines de semana, ya que en el apartado anterior se ha comprobado que ambos tipos de día presentan dinámicas distintas tanto en llamadas como en actividad futbolística.

Además, en las hipótesis centradas en resultados deportivos concretos (victorias, derrotas, remontadas, derbis o eliminaciones) se filtra únicamente por días con actividad futbolística registrada. Esta decisión permite que la comparación sea más coherente, ya que evita clasificar como “sin derrota” o “sin victoria” a días en los que directamente no hubo actividad futbolística registrada. De este modo, el análisis se centra en comparar distintos tipos de días con fútbol, en lugar de mezclar días con fútbol y días sin fútbol.

5.2.1 Hipótesis 1: los días con actividad futbolística registrada se asocian con un mayor número de llamadas al 016

La primera hipótesis plantea que los días con actividad futbolística registrada pueden presentar un mayor volumen medio de llamadas al 016 que los días sin actividad futbolística. Esta hipótesis funciona como aproximación inicial y general, antes de distinguir por tipo de resultado o intensidad deportiva.



El gráfico muestra que las llamadas medias son ligeramente superiores en los días con actividad futbolística registrada, tanto entre semana como en fin de semana. En los días laborables, la media de llamadas pasa de 6,6 en días sin actividad futbolística a 6,8 en días con actividad. En fines de semana, la diferencia también es positiva, pasando de 5,2 a 5,7 llamadas medias.

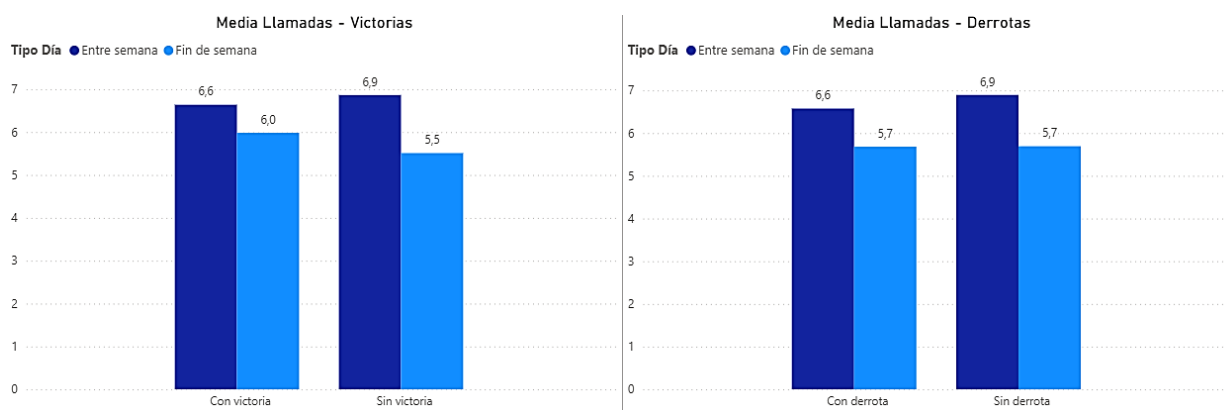
Este resultado sugiere una primera asociación positiva entre actividad futbolística registrada y llamadas al 016. No obstante, la diferencia es moderada, especialmente entre semana. Además, esta primera comparación no distingue todavía entre resultados favorables y desfavorables, por lo que no permite identificar si la posible relación procede del hecho de que haya fútbol en general o de determinados eventos deportivos concretos.

Aun así, el resultado es relevante como punto de partida: incluso después de diferenciar entre “entre semana” y “fin de semana”, los días con actividad futbolística muestran un volumen

medio de llamadas algo superior al de los días sin actividad. En los siguientes apartados se analiza si esta diferencia se concentra en determinados tipos de resultado o contextos deportivos.

5.2.2 Hipótesis 2: las derrotas se asocian con un mayor número de llamadas

La segunda hipótesis se centra en el resultado deportivo. A partir de la literatura previa, podría esperarse que las derrotas generen mayor frustración y por tanto, se asocien con un aumento en las llamadas al 016. Para analizarlo, se comparan las llamadas medias en días con y sin derrotas, filtrando únicamente por días con actividad futbolística registrada. Como contraste adicional también se representa la comparación entre días con y sin victorias.



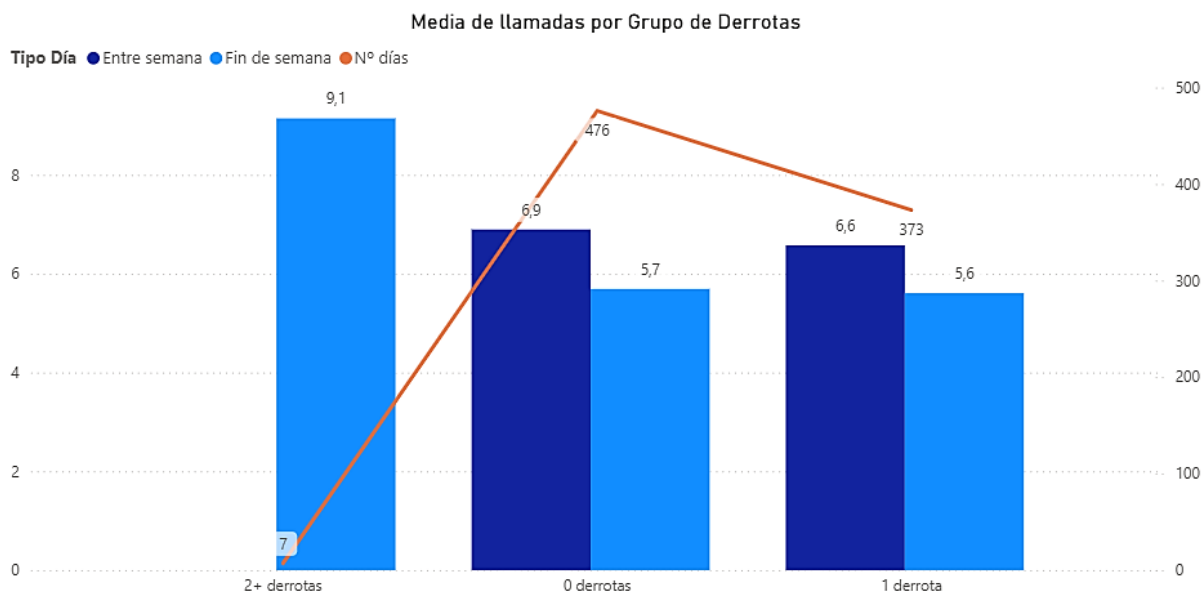
Los resultados no apoyan de forma directa la hipótesis de que las derrotas, consideradas de forma general, se asocien con más llamadas. En los días con actividad futbolística, las llamadas medias son muy similares entre días con derrota y días sin derrota. Entre semana, los días con derrota registran una media de 6,6 llamadas, frente a 6,9 en días sin derrota. En fin de semana, ambos grupos se sitúan prácticamente en el mismo nivel, con una media de 5,7 llamadas.

Este resultado obliga a matizar la hipótesis inicial. No parece suficiente distinguir únicamente entre derrota y no derrota. Es posible que el efecto emocional no dependa de cualquier derrota, sino de derrotas concretas con mayor carga emocional, como derrotas en descuento, remontadas sufridas, derbis o partidos de tensión alta.

El gráfico complementario de victorias tampoco muestra una asociación clara. En los días con victoria, las llamadas medias son de 6,6 entre semana y 6,0 en fin de semana, mientras que en los días sin victoria son de 6,9 y 5,5 respectivamente. Por tanto, las victorias tampoco presentan un patrón homogéneo: se asocian con menos llamadas entre semana, pero con más llamadas en fin de semana. Este resultado refuerza la idea de que el análisis debe avanzar hacia variables más específicas que capten mejor la intensidad emocional del evento deportivo.

5.2.3 Hipótesis 3: a mayor número de derrotas en la provincia, mayor número de llamadas

La tercera hipótesis plantea que el número de derrotas acumuladas en una provincia durante un mismo día podría estar relacionado con un mayor volumen de llamadas. Esta hipótesis es especialmente relevante en provincias con más de un equipo de referencia como Sevilla, donde puede coincidir actividad del Sevilla FC y del Real Betis.



El gráfico muestra que los días con dos o más derrotas presentan una media de llamadas elevada en fin de semana, alcanzando 9,1 llamadas. Sin embargo, este resultado debe interpretarse con mucha prudencia, ya que el número de días con dos o más derrotas es muy reducido: solo 7 observaciones. Por tanto, aunque el valor medio es alto, no puede considerarse evidencia robusta por sí mismo.

En cambio, al comparar los días con cero derrotas y una derrota, no se observa una relación creciente. Entre semana, los días con cero derrotas presentan una media de 6,9 llamadas, frente a 6,6 en días con una derrota. En fin de semana, las medias son prácticamente iguales: 5,7 en días con cero derrotas y 5,6 en días con una derrota.

Por tanto, esta hipótesis no encuentra un respaldo claro en el análisis descriptivo general. El dato de los días con dos o más derrotas resulta llamativo, pero el bajo número de observaciones impide extraer una conclusión firme. Este resultado muestra la importancia de incluir el número de días en los gráficos: sin esa referencia, podría interpretarse erróneamente que existe una relación fuerte entre acumulación de derrotas y llamadas, cuando en realidad el resultado depende de muy pocos casos.

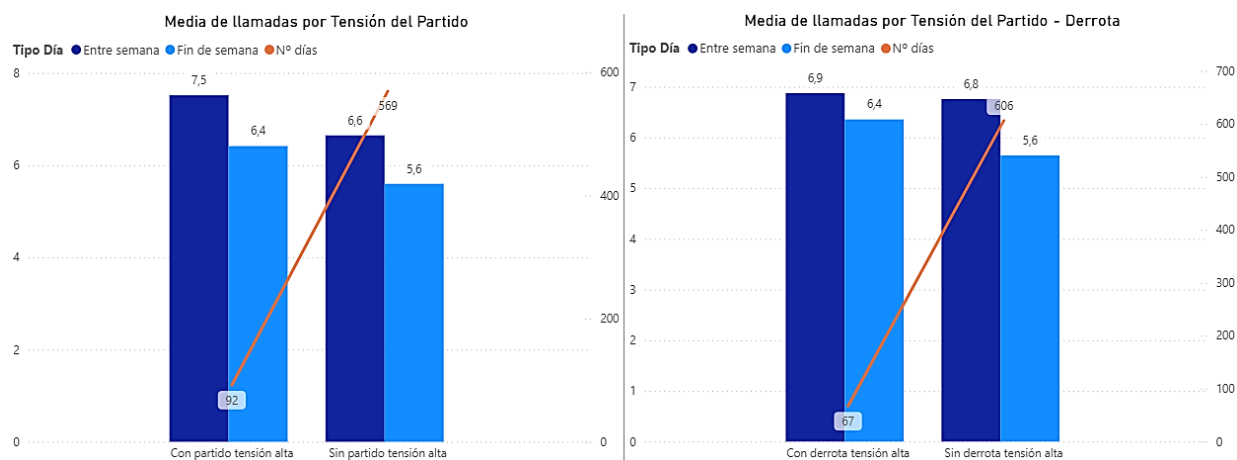
5.2.4 Hipótesis 4: los partidos de tensión alta se asocian con más llamadas

La cuarta hipótesis se basa en la idea de que no todos los partidos tienen la misma carga emocional. Algunos encuentros pueden generar una mayor tensión por su contexto competitivo o simbólico: derbis, eliminaciones o partidos en las últimas jornadas de liga. Por ello, se analiza primero la existencia de partidos de tensión alta y, posteriormente, las derrotas dentro de ese tipo de partidos.

El gráfico de partidos de tensión alta muestra una diferencia clara. En los días con partidos de tensión alta, las llamadas medias son superiores tanto entre semana como en fin de semana. Entre semana, la media alcanza 7,5 llamadas, frente a 6,6 en días sin partido de tensión alta. En fin de semana, la diferencia también es positiva: 6,4 frente a 5,6.

Este resultado ofrece un primer apoyo descriptivo a la hipótesis de que determinados contextos deportivos más intensos pueden asociarse con un mayor volumen de llamadas.

Además, el número de días con partido de tensión alta, aunque menor que el de días sin este tipo de partido, es suficiente para considerarlo un patrón relevante dentro del análisis descriptivo.



Cuando se analiza específicamente la derrota en partido de tensión alta, el resultado es más matizado. Entre semana, la diferencia entre días con y sin derrota de tensión alta es muy reducida: 6,9 llamadas frente a 6,8. Sin embargo, en fin de semana sí se observa una diferencia más clara: los días con derrota de tensión alta presentan una media de 6,4 llamadas, frente a 5,6 en días sin derrota de tensión alta.

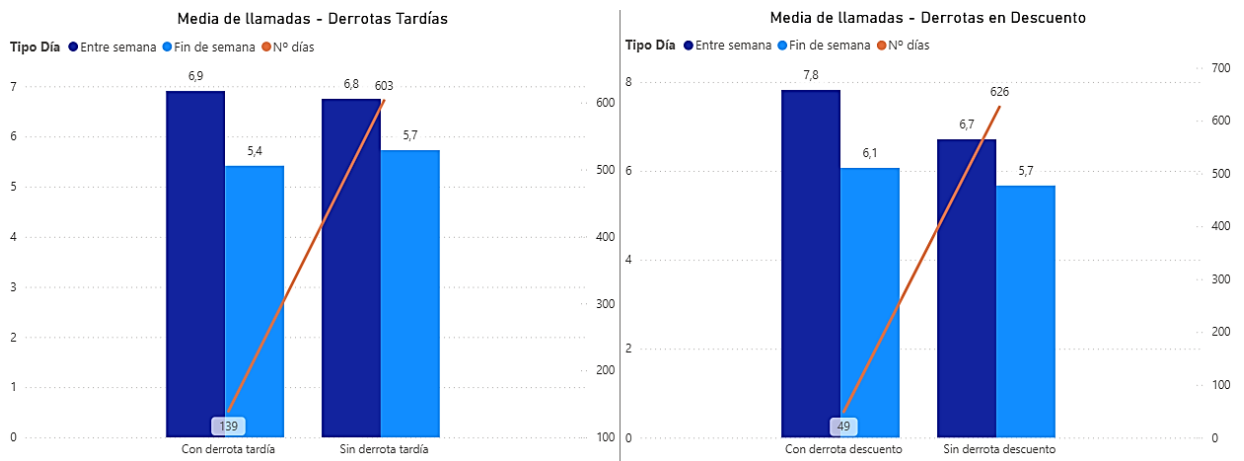
Este patrón sugiere que el contexto de tensión alta puede ser más relevante que la derrota por sí sola. En particular, el resultado parece más visible en fines de semana, que es cuando se concentra una parte importante de la actividad futbolística y cuando el componente social del partido puede ser mayor.

5.2.5 Hipótesis 5: las derrotas tardías y en descuento se asocian con más llamadas

La quinta hipótesis plantea que las derrotas producidas en los minutos finales pueden generar una frustración mayor que una derrota ordinaria. Para analizarlo, se distinguen dos variables: derrota tardía, asociada a goles encajados a partir del minuto 80, y derrota en descuento, asociada a goles encajados a partir del minuto 90 en partidos de margen estrecho.

El gráfico de derrotas tardías no muestra una asociación positiva clara con las llamadas. En los días con derrota tardía, la media de llamadas es de 6,9 entre semana y 5,4 en fin de semana. En los días sin derrota tardía, las medias son de 6,8 y 5,7 respectivamente. Por tanto, esta variable, definida de forma amplia a partir del minuto 80, no parece captar por sí sola un patrón de aumento en las llamadas.

Este resultado puede deberse a que no todos los goles tardíos tienen la misma carga emocional. Un gol recibido en el minuto 82 puede no tener el mismo efecto que un gol recibido en el descuento, especialmente si el partido ya estaba prácticamente decidido. Por ello, resulta útil analizar por separado la variable de derrota en descuento.

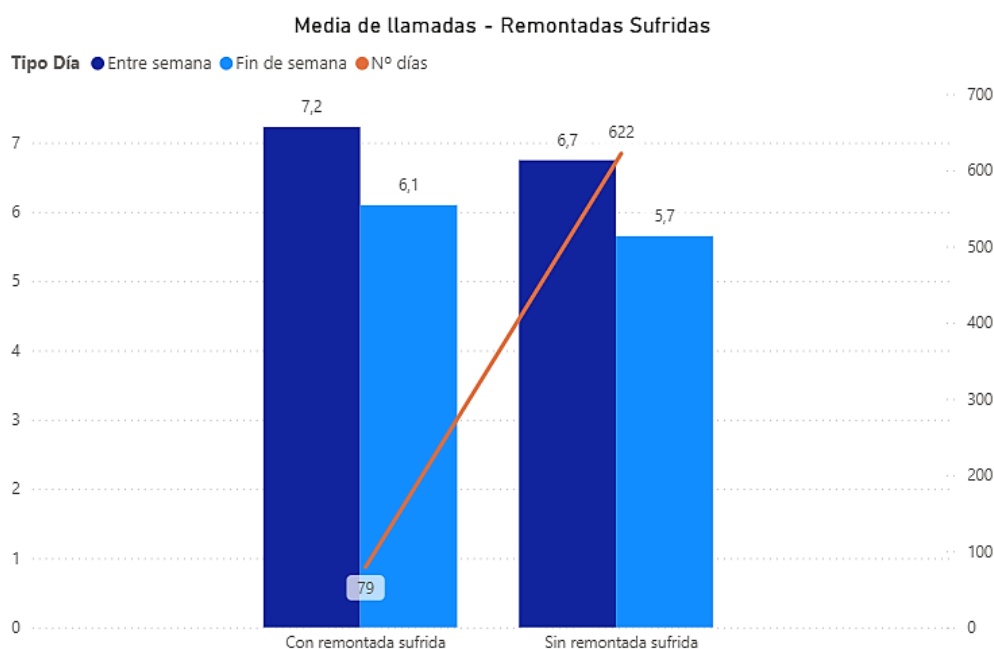


El gráfico de derrotas en descuento muestra un patrón más claro. En los días con derrota en descuento, las llamadas medias son superiores tanto entre semana como en fin de semana. Entre semana, la media alcanza 7,8 llamadas, frente a 6,7 en días sin derrota en descuento. En fin de semana, la diferencia también es positiva: 6,1 frente a 5,7.

Este resultado encaja mejor con la hipótesis emocional planteada. Perder en el descuento puede generar una frustración especialmente intensa, al producirse en el tramo final del partido y reducir la posibilidad de reacción. Sin embargo, el número de días con derrota en descuento es reducido, con 49 observaciones, por lo que el resultado debe interpretarse como indicio descriptivo y no como conclusión definitiva. Será necesario comprobar en el análisis estadístico si esta relación se mantiene al introducir controles temporales y territoriales.

5.2.6 Hipótesis 6: las remontadas sufridas se asocian con más llamadas

La sexta hipótesis considera que una derrota puede resultar más frustrante cuando el equipo llegó a ir ganando durante el partido. En este caso, no solo se produce una derrota, sino también una pérdida de ventaja, lo que puede intensificar la reacción emocional de los aficionados.



El gráfico muestra que los días con remontada sufrida presentan una media de llamadas

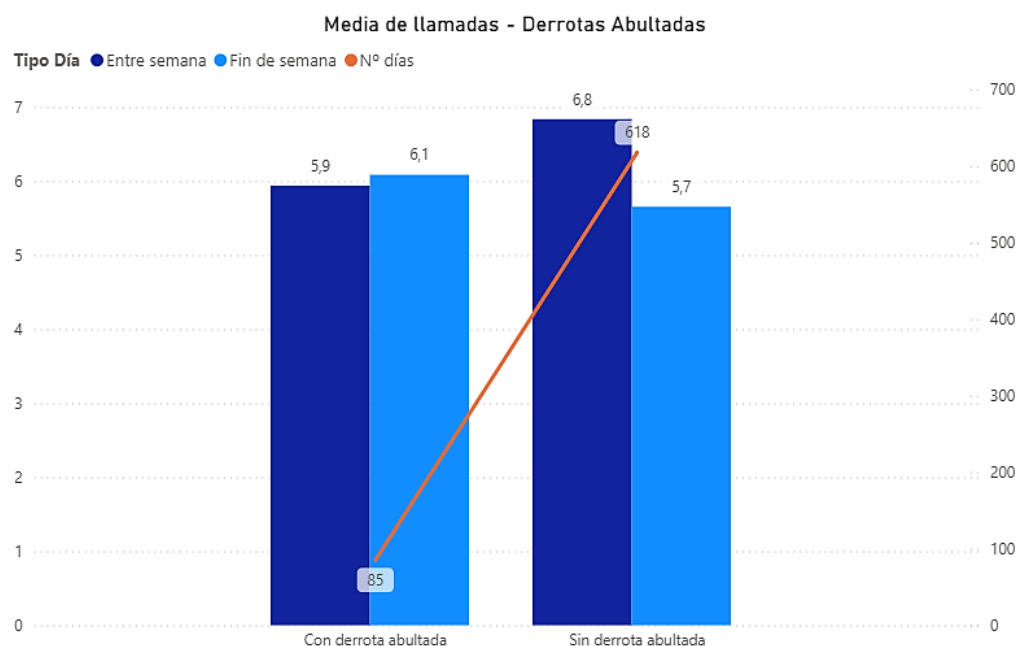
superior a los días sin remontada sufrida. Entre semana, la media es de 7,2 llamadas frente a 6,7. En fin de semana, la diferencia también es visible: 6,1 frente a 5,7.

Este resultado ofrece apoyo descriptivo a la hipótesis. Las remontadas sufridas parecen asociarse con un mayor volumen medio de llamadas, tanto en días laborables como en fines de semana. Además, aunque el número de días con remontada sufrida es menor que el de días sin este evento, no es tan reducido como en otras variables más excepcionales, con 79 observaciones.

Desde el punto de vista interpretativo, esta variable puede captar mejor la frustración deportiva que la derrota simple. No todas las derrotas son iguales: perder después de haber ido ganando puede percibirse como una oportunidad perdida y generar una reacción emocional más intensa.

5.2.7 Hipótesis 7: las derrotas abultadas se asocian con más llamadas

La séptima hipótesis plantea que las derrotas por una diferencia amplia de goles podrían estar asociadas a mayor tensión emocional al implicar una experiencia deportiva especialmente negativa o humillante para los aficionados.

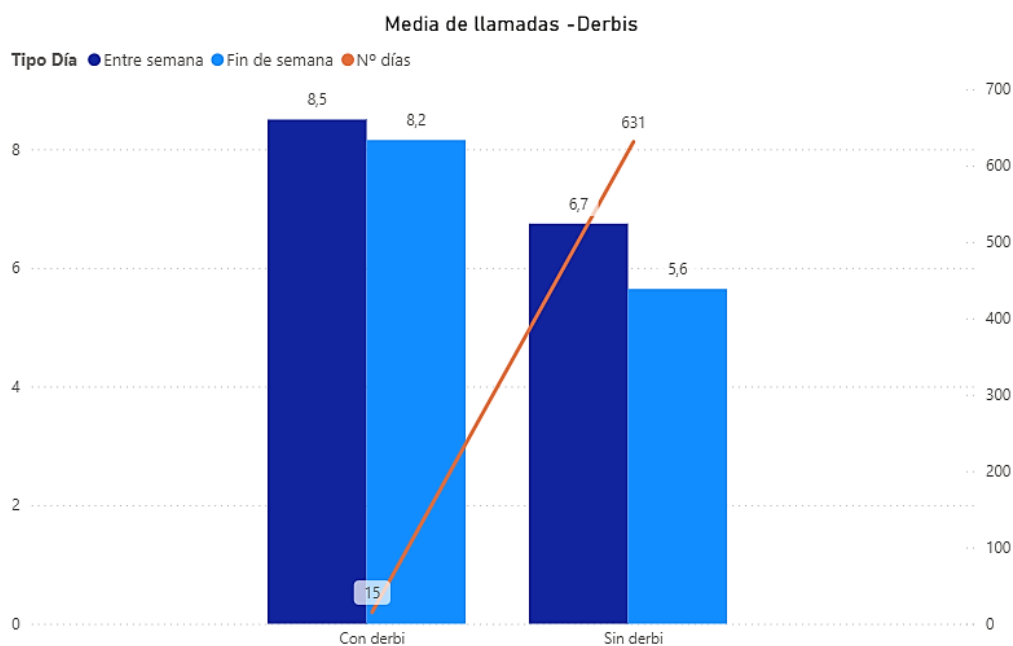


Los resultados son mixtos. Entre semana, los días con derrota abultada presentan una media de llamadas inferior a los días sin derrota abultada: 5,9 frente a 6,8. En fin de semana, ocurre lo contrario: los días con derrota abultada alcanzan 6,1 llamadas de media, frente a 5,7 en días sin derrota abultada.

Por tanto, la hipótesis no encuentra un respaldo claro. La diferencia positiva aparece únicamente en fin de semana y es moderada, mientras que entre semana el patrón va en sentido contrario. Esto puede deberse a que las derrotas abultadas no siempre generan la misma tensión que una derrota ajustada o inesperada. En algunos casos, si el resultado queda decidido pronto, el componente de incertidumbre y frustración final puede ser menor que en derrotas en descuento o remontadas sufridas.

5.2.8 Hipótesis 8: los derbis se asocian con más llamadas

La octava hipótesis se centra en los derbis. Estos partidos suelen tener una mayor carga emocional por la rivalidad territorial, la implicación social y la mayor atención mediática. Por ello, se analiza si los días con derbi presentan un volumen medio de llamadas superior al resto de días con actividad futbolística.

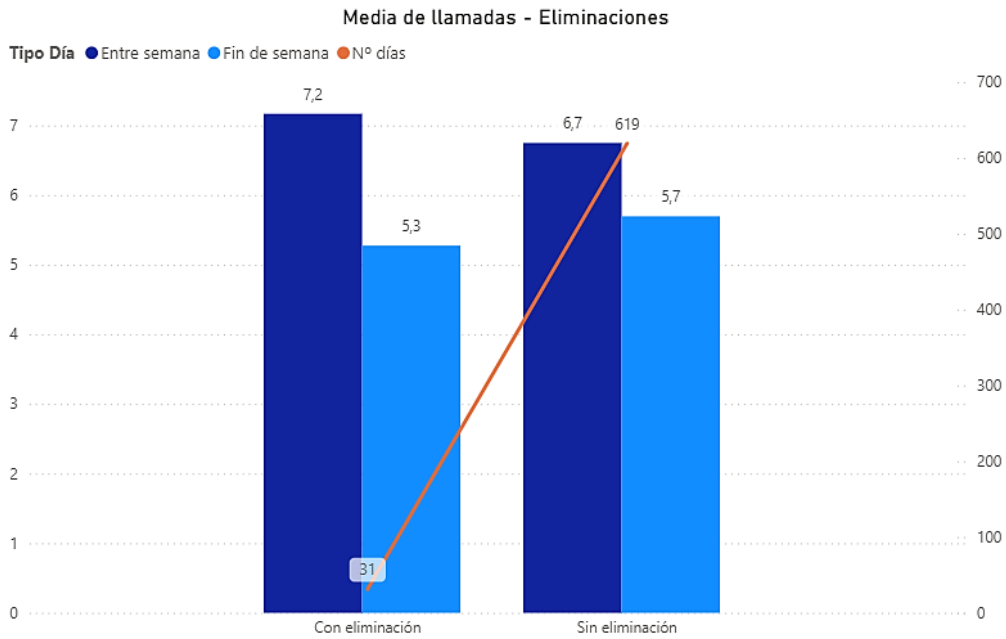


El gráfico muestra un resultado muy llamativo. Los días con derbi presentan una media de llamadas superior tanto entre semana como en fin de semana. Entre semana, la media alcanza 8,5 llamadas, frente a 6,7 en días sin derbi. En fin de semana, la diferencia también es clara: 8,2 frente a 5,6.

Este resultado apunta a que los derbis podrían estar asociados a un contexto emocional especialmente intenso. La diferencia es especialmente relevante porque aparece en ambos tipos de día y no solo en fines de semana. Sin embargo, la interpretación debe ser prudente porque el número de días con derbi es muy reducido, con solo 15 observaciones. Esto implica que la media puede estar muy influida por casos concretos. Aun así, el resultado es suficientemente relevante como para mantener esta variable en el análisis estadístico posterior, especialmente como parte de los indicadores de partidos de tensión alta.

5.2.9 Hipótesis 9: las eliminaciones se asocian con más llamadas

La novena hipótesis plantea que la eliminación de una competición puede tener un impacto emocional relevante, ya que supone el final de las expectativas deportivas en ese torneo. Esta variable se aplica principalmente a competiciones eliminatorias como el Playoff de Ascenso a Primera División, Copa del Rey o competiciones europeas.

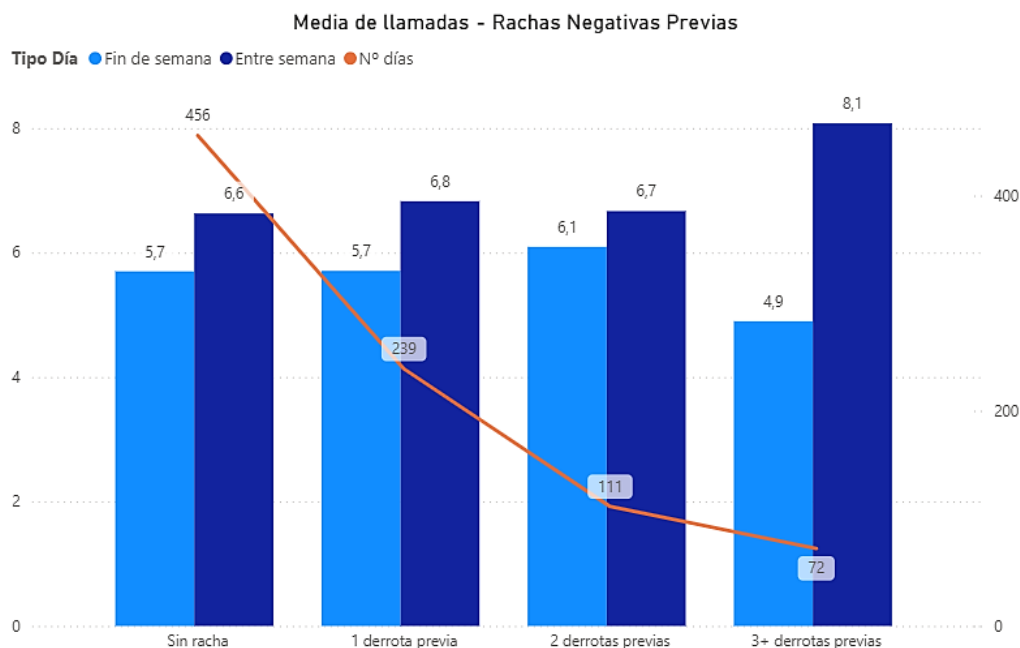


Los resultados vuelven a ser mixtos. Entre semana, los días con eliminación presentan una media de 7,2 llamadas, superior a los 6,7 de los días sin eliminación. En cambio, en fin de semana, los días con eliminación muestran una media inferior: 5,3 frente a 5,7.

Por tanto, no se observa una asociación positiva consistente entre eliminaciones y llamadas al 016. Además, el número de días con eliminación es limitado, con 31 observaciones. Esta variable puede seguir siendo útil como parte del indicador de partidos de tensión alta, pero por sí sola no parece mostrar un patrón descriptivo suficientemente claro.

5.2.10 Hipótesis 10: las rachas negativas previas se asocian con más llamadas

La décima hipótesis analiza si la acumulación de malos resultados previos puede generar un contexto de frustración deportiva. Para ello, se compara la media de llamadas según la racha negativa previa del equipo.



El gráfico no muestra una relación lineal clara entre racha negativa y llamadas. Entre semana, las llamadas medias se sitúan en 6,6 en días sin racha, 6,8 con una derrota previa, 6,7 con dos

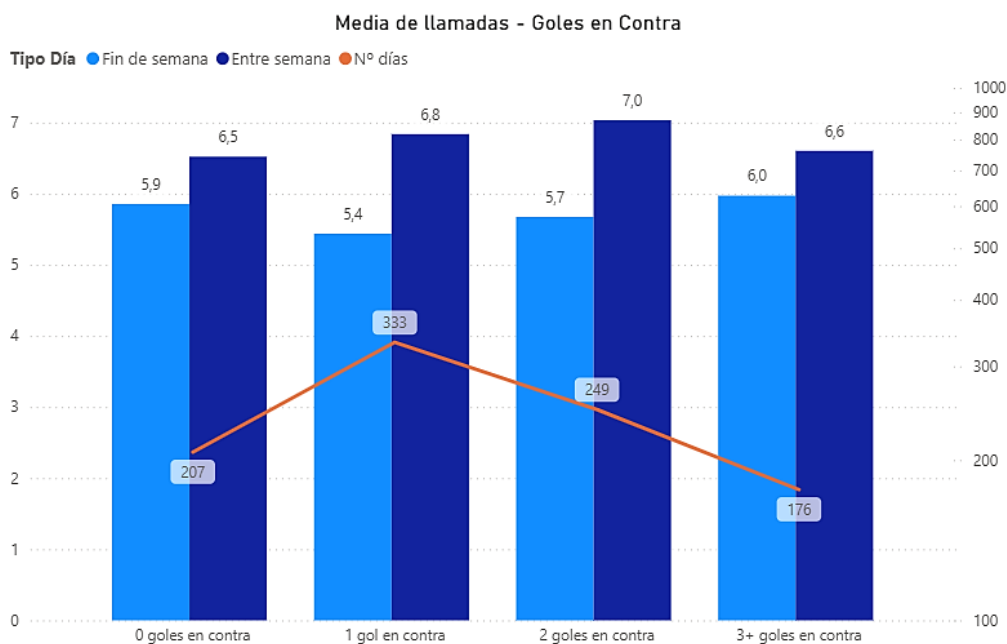
derrotas previas y 8,1 con tres o más derrotas previas. Este último valor es elevado, aunque debe interpretarse con cautela porque el número de observaciones es menor.

En fin de semana, el patrón tampoco es creciente. La media es de 5,7 en días sin racha, 5,7 con una derrota previa, 6,1 con dos derrotas previas y 4,9 con tres o más derrotas previas.

Estos resultados no apoyan de forma clara la hipótesis de que las rachas negativas previas incrementen las llamadas. Una posible explicación es que la racha previa, por sí sola, no representa necesariamente un evento emocional inmediato. Puede contribuir al contexto general de frustración, pero el efecto más visible podría depender del resultado concreto del partido analizado o de la importancia competitiva del encuentro.

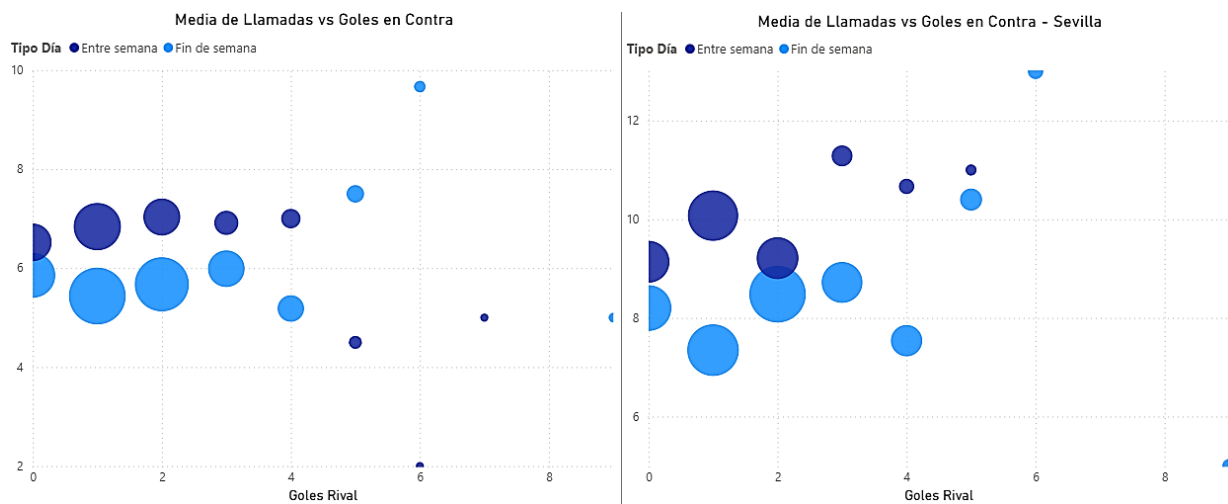
5.2.11 Hipótesis 11: los goles en contra se asocian con más llamadas

Además del resultado final, se analiza si el número de goles encajados por el equipo de referencia se relaciona con el volumen de llamadas. Esta variable permite estudiar una relación más gradual que la simple distinción entre victoria, empate o derrota.



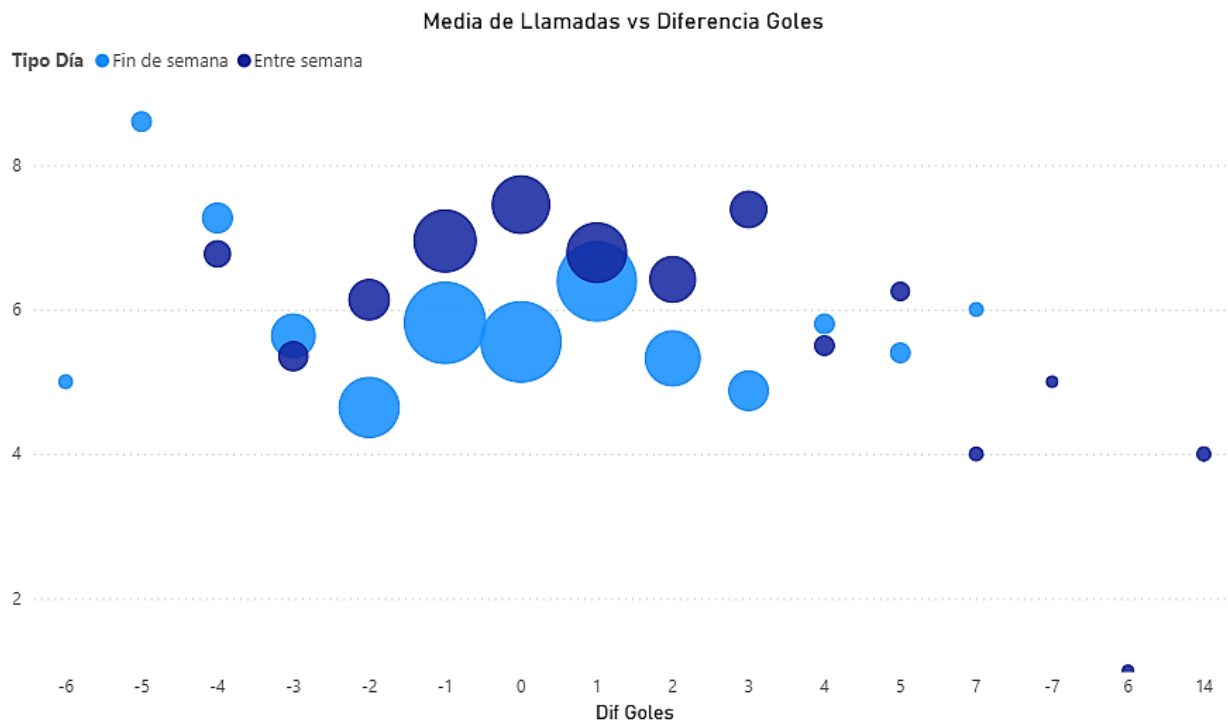
El gráfico de medias por grupos de goles en contra no muestra una relación lineal clara. Entre semana, la media de llamadas aumenta ligeramente al pasar de 0 a 2 goles en contra: 6,5 llamadas con 0 goles en contra, 6,8 con 1 gol y 7,0 con 2 goles. Sin embargo, en el grupo de 3 o más goles en contra la media baja a 6,6. En fin de semana, el patrón también es irregular: 5,9 llamadas con 0 goles en contra, 5,4 con 1 gol, 5,7 con 2 goles y 6,0 con 3 o más goles.

Por tanto, el análisis agregado no permite afirmar que a mayor número de goles en contra haya necesariamente más llamadas. El patrón parece depender del tipo de día y posiblemente de la provincia. Para profundizar en esta relación, se utilizan gráficos de dispersión entre goles del rival y llamadas medias.



La dispersión general entre goles del rival y llamadas medias no muestra una relación clara. Los puntos aparecen relativamente dispersos y no se observa una tendencia lineal evidente. Esto sugiere que el número de goles encajados, considerado de forma agregada para todas las provincias, no parece explicar por sí solo las variaciones en las llamadas.

Sin embargo, al observar específicamente la provincia de Sevilla, el patrón parece algo más definido. En este caso, algunos puntos con mayor número de goles en contra se asocian con medias de llamadas más elevadas. Esto no implica una relación causal ni necesariamente robusta, pero puede indicar que la asociación entre fútbol y llamadas es más visible en territorios con mayor volumen de llamadas y mayor exposición futbolística. En cualquier caso, esta posibilidad deberá contrastarse posteriormente mediante modelos que controlen por provincia y por factores temporales.



También se analiza la relación entre la diferencia de goles y la media de llamadas. En este caso, el gráfico no muestra una correlación clara. Las llamadas medias no aumentan de forma sistemática cuando la diferencia de goles es más negativa, ni disminuyen claramente cuando el resultado deportivo es más favorable. Este resultado refuerza la idea de que las variables

deportivas agregadas de forma simple pueden no captar completamente la dimensión emocional del partido. Algunas variables más específicas, como remontadas sufridas, derbis o derrotas en descuento, parecen ofrecer una lectura más útil que la diferencia de goles tomada de forma aislada.

5.2.12 Síntesis del contraste de hipótesis diarias

El análisis descriptivo diario ofrece resultados heterogéneos. La hipótesis más general, según la cual los días con actividad futbolística registrada presentan más llamadas, encuentra un primer apoyo descriptivo, aunque con diferencias moderadas. Tanto entre semana como en fin de semana, las llamadas medias son algo superiores cuando hay actividad futbolística registrada. Sin embargo, al entrar en el resultado deportivo, la relación no es tan sencilla.

Las derrotas consideradas de forma general no se asocian con un aumento claro de llamadas. En los días con actividad futbolística, las llamadas medias son muy similares entre días con derrota y días sin derrota. Las victorias tampoco muestran un patrón homogéneo. Este resultado obliga a rechazar una interpretación demasiado simple basada únicamente en victoria, empate o derrota.

En cambio, algunas variables más específicas sí muestran patrones más coherentes con la hipótesis emocional del trabajo. Entre las variables que muestran resultados más relevantes destacan los partidos de tensión alta, las derrotas en descuento, las remontadas sufridas y los derbis. Los partidos de tensión alta presentan medias superiores tanto entre semana como en fin de semana. Las derrotas en descuento y las remontadas sufridas también muestran diferencias positivas. Los derbis presentan uno de los patrones más llamativos, aunque basado en un número muy reducido de observaciones.

Por el contrario, otras variables como las derrotas tardías, las derrotas abultadas, las eliminaciones o las rachas negativas previas no muestran una relación descriptiva consistente. En algunos casos, los resultados son mixtos según el tipo de día. En otros, la diferencia es reducida o depende de pocas observaciones.

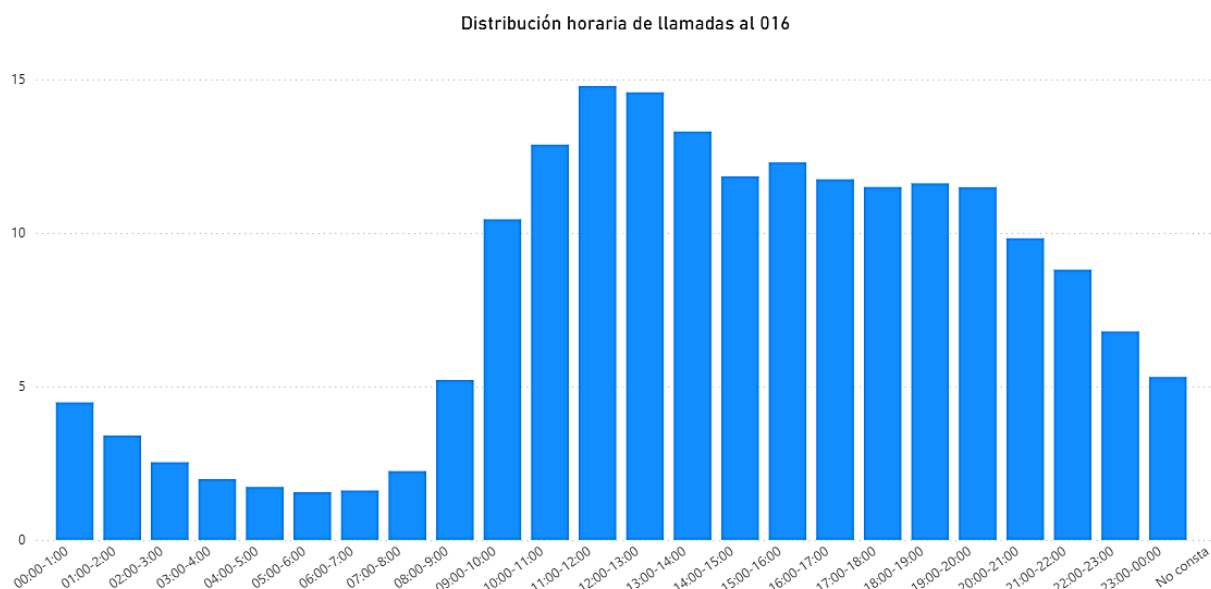
En conjunto, estos resultados sugieren que el posible vínculo entre fútbol y llamadas al 016 no parece depender simplemente de que un equipo pierda, sino del tipo de evento deportivo y del contexto en el que se produce. Los eventos más ajustados, emocionalmente intensos o socialmente relevantes parecen ofrecer una explicación más plausible que las variables deportivas generales. Esta conclusión orienta el análisis estadístico posterior, en el que se contrastarán estas relaciones incorporando controles por provincia, tipo de día y periodo temporal.

5.3 Análisis exploratorio por franjas horarias

Además del análisis diario, se realiza una exploración de los patrones horarios utilizando la base de llamadas por franja. Como se ha explicado previamente, esta base recoge llamadas acumuladas por provincia, mes y franja horaria, por lo que no permite vincular una llamada concreta con un partido concreto. Por tanto, los resultados de esta sección deben interpretarse como descriptivos y exploratorios.

El objetivo de este bloque no es estimar un efecto horario preciso de los partidos, sino comprobar si los patrones agregados de llamadas al 016 se aproximan a la distribución temporal de la actividad futbolística. Para ello se comparan las llamadas medias por franja con tres medidas distintas de exposición deportiva: partidos en juego durante la franja, partidos asignados a la franja posterior a su finalización y goles en contra registrados durante la franja.

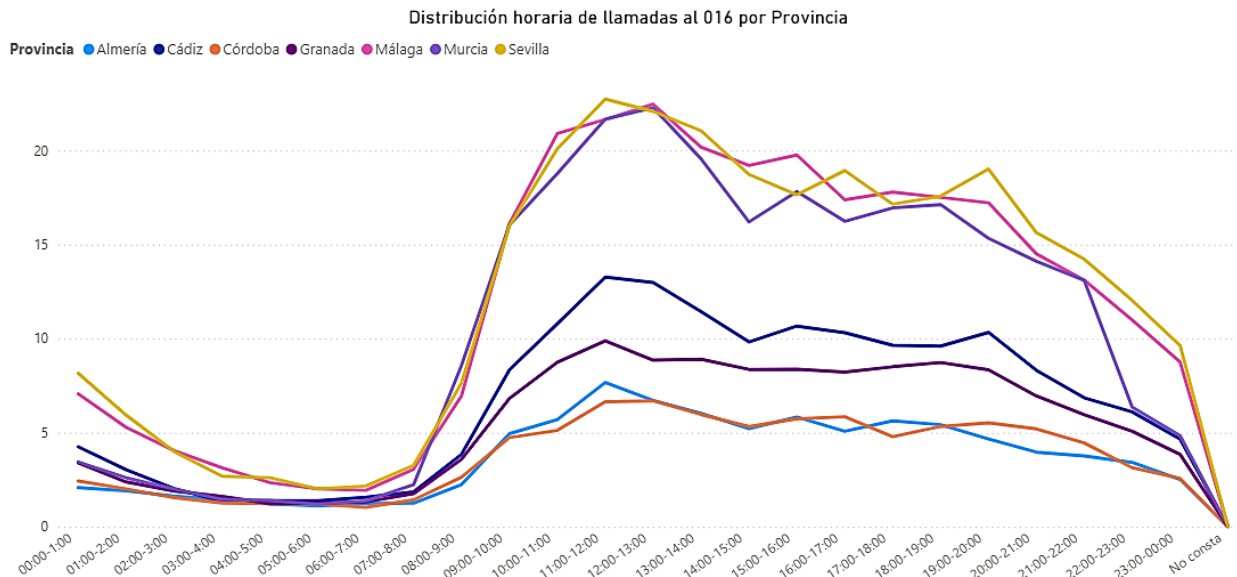
5.3.1 Distribución horaria general de las llamadas



La distribución horaria de llamadas al 016 muestra un patrón muy claro. Las llamadas son reducidas durante la madrugada y primeras horas de la mañana. Entre las 00:00 y las 08:00, el volumen medio se mantiene en niveles bajos, con mínimos en torno a las franjas de 05:00-06:00 y 06:00-07:00. A partir de las 08:00-09:00 comienza un incremento progresivo, que se acelera especialmente desde las 09:00-10:00.

Los valores máximos se concentran entre media mañana y primeras horas de la tarde. Las franjas de 11:00-12:00, 12:00-13:00 y 13:00-14:00 presentan los niveles más elevados de llamadas, con medias cercanas a 15 llamadas. A partir de ese momento, el volumen desciende de forma gradual, aunque se mantiene en niveles relativamente altos durante la tarde. Desde las 20:00-21:00 comienza una caída más clara, que se acentúa en las últimas horas del día.

Este patrón es coherente con una mayor utilización del servicio durante horarios diurnos y laborales. Además, resulta especialmente relevante para interpretar la comparación con la actividad futbolística, ya que la mayor parte de los partidos se concentra en franjas de tarde y noche, no en las horas de mayor volumen de llamadas.

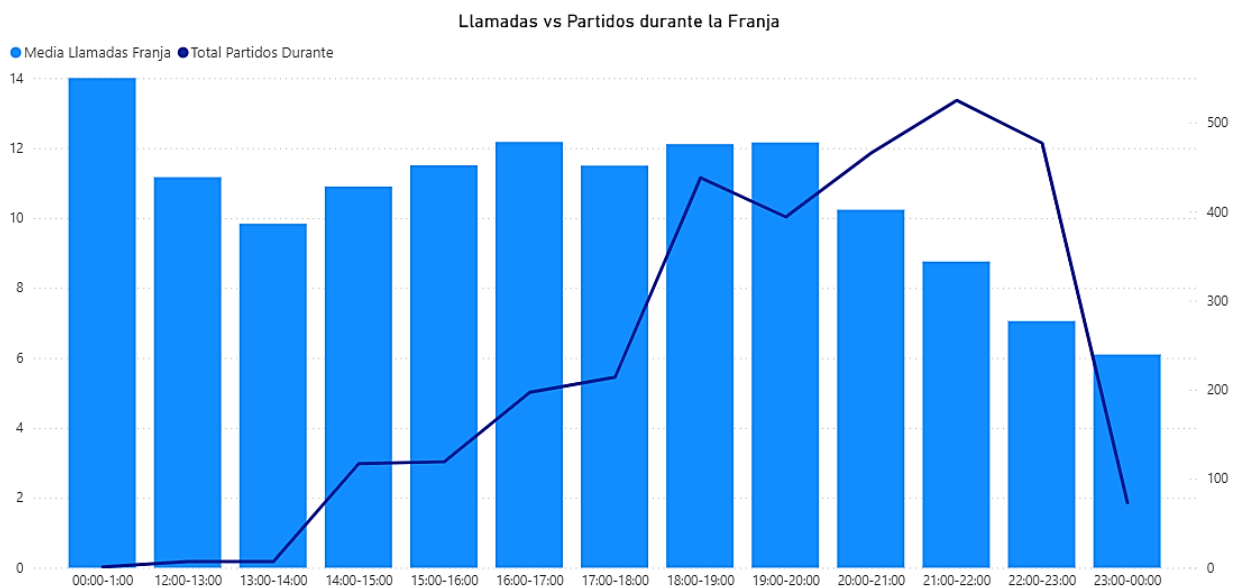


La distribución horaria por provincia mantiene una forma general similar en todos los territorios, aunque con diferencias de nivel. Sevilla, Málaga y Murcia presentan los valores más elevados en la mayor parte de las franjas, especialmente entre media mañana y primeras horas de la tarde. Cádiz y Granada se sitúan en un nivel intermedio, mientras que Almería y Córdoba presentan medias más reducidas.

A pesar de estas diferencias territoriales, el patrón horario es común: baja actividad durante la madrugada, crecimiento rápido a partir de la mañana, máximo alrededor del mediodía y descenso progresivo durante la tarde y la noche. Este resultado sugiere que la distribución temporal de las llamadas responde a una dinámica general del servicio, más que a un patrón específico de una provincia concreta.

También resulta relevante que las diferencias entre provincias sean más visibles en las franjas de mayor actividad. Durante la madrugada, las provincias se encuentran más próximas entre sí, mientras que entre las 10:00 y las 14:00 las diferencias de nivel se amplían claramente. Esto indica que las provincias con mayor volumen total de llamadas no solo llaman más en conjunto, sino que concentran especialmente esa diferencia en las horas centrales del día.

5.3.2 Llamadas y partidos durante la franja



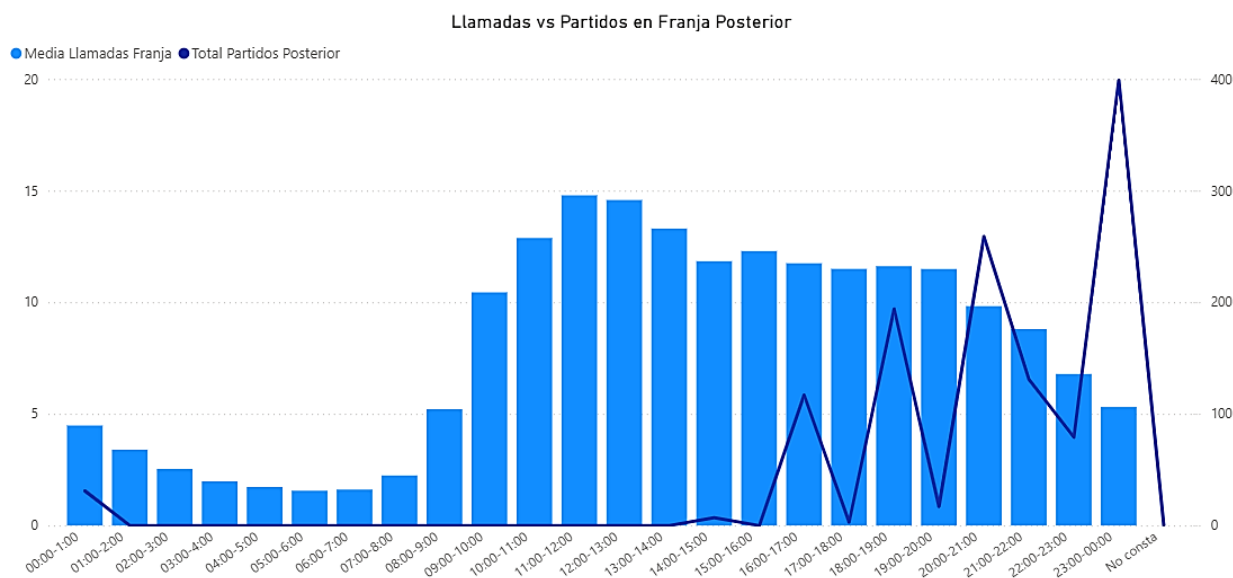
El gráfico que compara llamadas medias por franja y partidos en juego durante esa franja muestra una diferencia temporal clara entre ambas series. Las llamadas alcanzan sus valores más altos durante la mañana y el inicio de la tarde, mientras que los partidos en juego se concentran principalmente en franjas de tarde y noche.

La actividad futbolística durante la franja comienza a ganar peso a partir de las 14:00-15:00 y aumenta de forma clara desde las 18:00-19:00. Los máximos de partidos durante la franja se sitúan entre las 20:00 y las 23:00, especialmente en las franjas de 21:00-22:00 y 22:00-23:00. En cambio, en esas horas las llamadas ya se encuentran en descenso respecto a sus máximos del mediodía.

Esta falta de coincidencia temporal sugiere que, a nivel agregado mensual, no parece existir una relación directa entre las franjas con más partidos en juego y las franjas con más llamadas. De hecho, las horas de mayor actividad futbolística coinciden con un tramo en el que las llamadas ya han comenzado a reducirse.

Este resultado no permite descartar efectos concretos durante partidos específicos, ya que no se dispone de llamadas por hora y día. Sin embargo, sí indica que el patrón horario general de las llamadas al 016 no parece estar determinado por la distribución horaria de los partidos.

5.3.3 Llamadas y franja posterior al partido



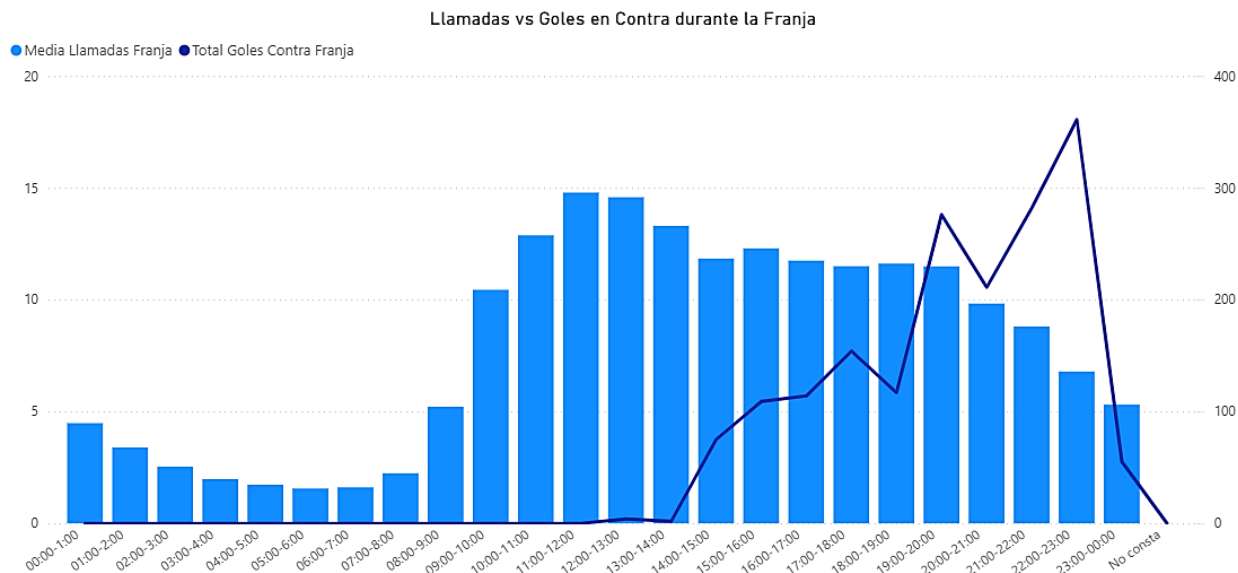
La comparación entre llamadas y partidos asignados a la franja posterior al encuentro ofrece una conclusión similar. Las llamadas se concentran principalmente en horario diurno, con máximos entre las 11:00 y las 14:00, mientras que las franjas posteriores a los partidos se concentran en horarios de tarde-noche y noche.

La serie de partidos en franja posterior presenta picos especialmente visibles en las franjas de 19:00-20:00, 21:00-22:00 y 23:00-00:00. Esta distribución es coherente con los horarios habituales de finalización de los encuentros. Sin embargo, esas franjas no coinciden con los valores máximos de llamadas, que se producen varias horas antes.

Esta diferencia temporal limita la posibilidad de observar una asociación horaria clara en la base mensual por franjas. Aunque la literatura previa apunta a que algunos efectos pueden

producirse después del partido, los datos disponibles en este trabajo no permiten comprobarlo con precisión horaria diaria. A nivel agregado, las franjas posteriores a los partidos no coinciden con las franjas de mayor volumen de llamadas.

5.3.4 Llamadas y goles en contra durante la franja



El gráfico de llamadas frente a goles en contra durante la franja refuerza la misma idea. Los goles en contra se concentran en las franjas de tarde y noche, coincidiendo con los horarios habituales de los partidos. La serie de goles en contra aumenta a partir de media tarde y alcanza sus valores más elevados en horario nocturno, especialmente entre las 20:00 y las 23:00.

En cambio, las llamadas presentan sus máximos entre media mañana y primeras horas de la tarde. Por tanto, no se observa una coincidencia clara entre los momentos horarios con más goles en contra y los momentos horarios con más llamadas. De hecho, cuando los goles en contra alcanzan su mayor presencia, las llamadas ya han iniciado una tendencia descendente.

Este resultado debe interpretarse con cautela, ya que el dato de llamadas no está disponible por día y hora, sino agregado mensualmente. Aun así, desde un punto de vista descriptivo, no parece que el patrón horario general de las llamadas esté alineado con el patrón horario de los goles o de los partidos.

5.3.5 Síntesis del análisis por franjas

El análisis por franjas horarias muestra que las llamadas al 016 siguen un patrón diario agregado muy marcado: son bajas durante la madrugada, aumentan durante la mañana, alcanzan su máximo alrededor del mediodía y descienden progresivamente durante la tarde y la noche. Este patrón se reproduce en todas las provincias, aunque con distintos niveles absolutos.

Por el contrario, la actividad futbolística se concentra principalmente en horario de tarde y noche. Tanto los partidos durante la franja como las franjas posteriores al partido y los goles en contra muestran mayor presencia en las horas finales del día. Esta diferencia entre el horario habitual de llamadas y el horario habitual de los partidos dificulta encontrar una

asociación descriptiva clara en el análisis horario agregado.

En consecuencia, el análisis por franjas no ofrece evidencia visual fuerte de una relación entre llamadas y actividad futbolística en una misma franja horaria. Su principal aportación es mostrar que los patrones horarios de ambas variables son distintos. Por ello, el análisis principal del trabajo debe apoyarse en la base diaria por provincia, mientras que la base por franjas se mantiene como complemento exploratorio.

5.4 Conclusión del análisis descriptivo

El análisis descriptivo realizado en Power BI permite extraer varias conclusiones relevantes. En primer lugar, las llamadas al 016 presentan una clara heterogeneidad territorial. Sevilla, Málaga y Murcia concentran los mayores volúmenes de llamadas dentro del ámbito analizado, mientras que Córdoba y Almería presentan los niveles más reducidos.

En segundo lugar, existe una diferencia muy marcada entre días laborables y fines de semana. Las llamadas son más elevadas entre semana, especialmente al comienzo de la semana, mientras que la actividad futbolística se concentra principalmente en sábado y domingo. Esta diferencia es fundamental para el análisis, ya que impide interpretar de forma directa la comparación entre días con y sin fútbol sin tener en cuenta el tipo de día.

En tercer lugar, los días con actividad futbolística registrada presentan llamadas medias ligeramente superiores a los días sin actividad, tanto entre semana como en fin de semana. No obstante, al analizar el resultado deportivo, las derrotas consideradas de forma general no muestran una asociación positiva clara con las llamadas. Las victorias tampoco presentan un patrón homogéneo. Esto sugiere que la relación entre fútbol y llamadas al 016 no puede reducirse a la simple existencia de una victoria o una derrota.

Las variables que parecen mostrar patrones descriptivos más relevantes son aquellas que captan situaciones deportivas más específicas o emocionalmente intensas: partidos de tensión alta, derrotas en descuento, remontadas sufridas y derbis. En cambio, otras variables como derrotas tardías, derrotas abultadas, eliminaciones o rachas negativas previas no presentan una relación consistente.

Por último, el análisis por franjas horarias muestra que las llamadas y la actividad futbolística tienen patrones horarios diferentes. Las llamadas se concentran principalmente en horario diurno, con máximos entre media mañana y primeras horas de la tarde, mientras que los partidos y goles se concentran en la tarde y noche. Por ello, el análisis horario debe interpretarse como exploratorio y no como evidencia directa de un efecto durante o inmediatamente después del partido.

En conjunto, los resultados descriptivos apuntan a que la posible asociación entre fútbol y llamadas al 016, si existe, estaría más relacionada con determinados contextos deportivos de alta intensidad que con la actividad futbolística o las derrotas consideradas de forma general. Esta conclusión sirve como base para el análisis estadístico posterior en RStudio, donde se evaluará si estas relaciones se mantienen al introducir controles por provincia, tipo de día y periodo temporal.

6. Análisis Estadístico y Econométrico –RStudio

6.1. Preparación de las bases para el análisis estadístico

Una vez finalizada la construcción del modelo de datos en Power BI, las tablas finales se exportaron a formato CSV para realizar el análisis estadístico en RStudio. El objetivo de esta fase es comprobar si las relaciones observadas en el análisis descriptivo se mantienen al introducir controles temporales y territoriales, así como evaluar la robustez de los resultados mediante distintas especificaciones.

El análisis se apoya en dos bases principales. La primera, Dataset_R_Diario, constituye la base central del estudio y trabaja con una unidad de observación provincia-fecha. La segunda, Dataset_R_Franjas, se utiliza para el análisis horario agregado y trabaja con una unidad de observación provincia-mes-franja horaria.

Resumen de bases preparadas			
base	observaciones	columnas	duplicados_clave
diario	13086	74	0
franjas	9841	57	0

La base diaria contiene 13.086 observaciones y 74 variables. La base por franjas contiene 9.841 observaciones y 57 variables. En ambos casos, la comprobación de claves no muestra duplicidades: la base diaria mantiene una única fila por provincia y fecha, y la base por franjas mantiene una única fila por provincia, mes y franja horaria. Esta comprobación es importante porque garantiza que las llamadas no se dupliquen artificialmente al cruzarlas con variables futbolísticas.

6.1.1. Base diaria provincia-fecha

En la base diaria la variable dependiente es el número de llamadas pertinentes al 016. Las variables explicativas principales son las variables deportivas construidas a partir del registro de partidos: actividad futbolística registrada, victorias, derrotas, empates, goles a favor, goles en contra, diferencia de goles, derrotas tardías, derrotas en descuento, remontadas sufridas, derbis, eliminaciones, partidos de tensión alta y derrotas en partidos de tensión alta, entre otras.

En las fechas sin actividad futbolística registrada las variables deportivas toman valor cero. Esta decisión permite mantener todas las observaciones de llamadas, tanto en días con fútbol como en días sin fútbol, y evita limitar el análisis únicamente a jornadas con partidos. No obstante, en algunos contrastes específicos, como la comparación entre días con y sin derrota, se filtra previamente por días con actividad futbolística registrada para que no se clasifiquen como “sin derrota” días en los que directamente no hubo fútbol.

También se incorporaron controles temporales y territoriales. En concreto, se incluyeron variables de provincia, mes-año y tipo de día. La inclusión de provincia es fundamental porque las llamadas presentan niveles muy distintos entre territorios. La inclusión de mes-año permite controlar variaciones temporales comunes, como estacionalidad, cambios de tendencia o diferencias entre años. Finalmente, la variable de fin de semana permite controlar la diferencia estructural entre días laborables y fines de semana.

6.1.2. Base mensual provincia-franja horaria

La base por franjas se utiliza para estudiar patrones horarios agregados. Cada observación representa una provincia, un mes y una franja horaria. La variable dependiente es el número de llamadas acumuladas en esa franja durante el mes. Las variables futbolísticas se agrupan en tres familias: partidos durante la franja, partidos asignados a la franja posterior a su finalización y goles registrados dentro de la franja.

Este análisis tiene una limitación importante: no permite saber si una llamada concreta se produjo el mismo día que un partido concreto ya que la información horaria del 016 está agregada mensualmente. Por ello, los resultados de franjas se interpretan únicamente como una exploración de patrones horarios generales y no como evidencia causal.

6.1.3. Tratamiento de variables y controles temporales

Antes de estimar los modelos, las variables numéricas se transformaron a formato adecuado y se crearon indicadores binarios para los principales eventos deportivos. Por ejemplo, hay_futbol toma valor 1 si en esa provincia y fecha existe actividad futbolística registrada, y 0 en caso contrario. De forma similar, se crearon variables como hay_derrota, hay_victoria, hay_derrota_descuento, hay_remontada_sufrida, hay_derbi y hay_partido_tension_alta.

Además se construyeron variables retardadas para analizar si algunos efectos podían aparecer al día siguiente del partido. En particular, se generaron retardos de actividad futbolística, derrota, derrota en descuento, remontada sufrida y partido de tensión alta. Esta decisión se apoya en la idea de que una posible reacción asociada al evento deportivo no tiene por qué producirse necesariamente el mismo día, sino que puede trasladarse parcialmente al día posterior.

6.2. Análisis exploratorio inicial

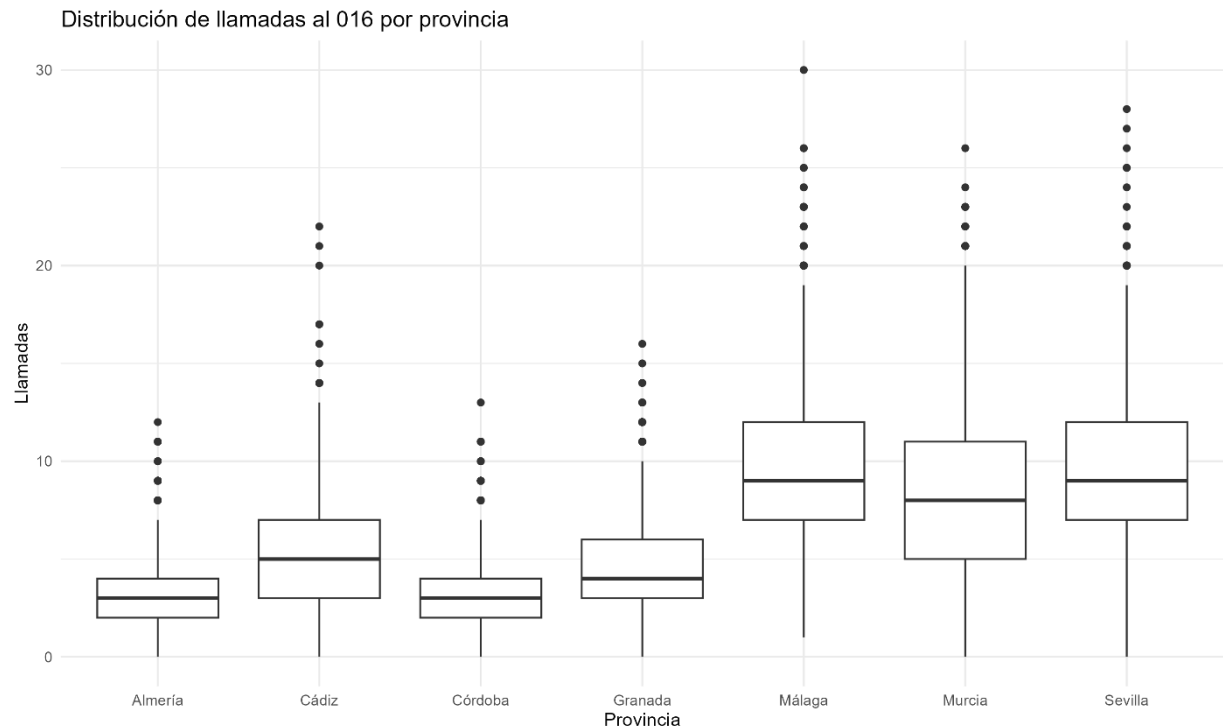
6.2.1. Estadísticos descriptivos de las variables principales

El primer paso del análisis estadístico consistió en calcular los estadísticos descriptivos de la variable dependiente y de las principales variables explicativas. La media general de llamadas en la base diaria es de 6,32 llamadas por provincia y día, con una mediana de 5 llamadas. La desviación típica es de 4,32, lo que refleja una dispersión considerable entre observaciones. El valor máximo observado es de 30 llamadas en una provincia y fecha concreta.

Estadísticos descriptivos generales de llamadas						
observaciones	media	mediana	sd	min	max	total
13.086	6,322099954	5	4,316742328	0	30	82.731

El análisis por provincia muestra diferencias territoriales relevantes. Sevilla presenta la media diaria más elevada, con 9,85 llamadas, seguida muy de cerca por Málaga, con 9,60. Murcia ocupa el tercer puesto, con 8,53 llamadas medias. A continuación se sitúan Cádiz, con 5,34; Granada, con 4,39; Córdoba, con 3,06; y Almería, con 3,01.

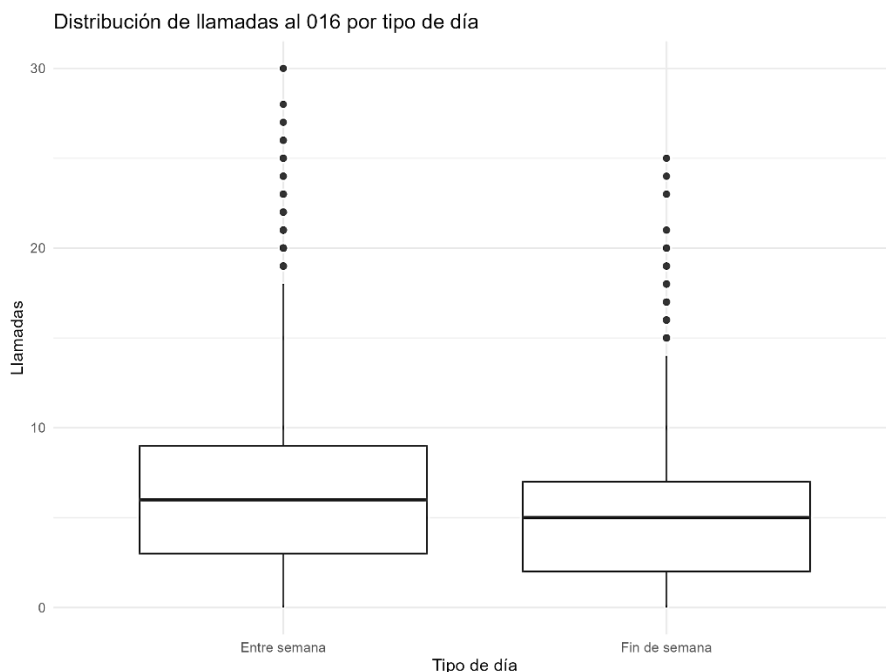
Estadísticos por provincia										
provincia	observaciones	media	mediana	sd	min	max	total	partidos	derrotas	victorias
Sevilla	1915	9,845430809	9	4,10839594	0	28	18854	384	119	178
Málaga	1913	9,602718244	9	4,133118729	1	30	18370	118	54	36
Murcia	1908	8,529874214	8	4,395932785	0	26	16275	165	88	59
Cádiz	1896	5,342299578	5	2,879354917	0	22	10129	162	77	46
Granada	1877	4,388918487	4	2,479348553	0	16	8238	174	79	61
Córdoba	1791	3,062534897	3	1,961215575	0	13	5485	42	15	15
Almería	1786	3,012318029	3	1,950612775	0	12	5380	177	69	70



El boxplot por provincia confirma visualmente esta heterogeneidad territorial. Sevilla, Málaga y Murcia presentan distribuciones más elevadas, mientras que Almería, Córdoba y Granada muestran valores más reducidos y concentrados. Cádiz se sitúa en una posición intermedia. Esta diferencia territorial justifica la inclusión de efectos de provincia en los modelos, ya que las llamadas no parten del mismo nivel en todos los territorios.

También se observan diferencias importantes por tipo de día. Entre semana, la media de llamadas es de 6,71, mientras que en fin de semana desciende a 5,34. Sin embargo, la actividad futbolística se comporta en sentido contrario: la media de partidos registrados es de 0,046 entre semana y de 0,213 en fin de semana. Esta diferencia confirma la necesidad de controlar por calendario semanal, ya que los días con más fútbol coinciden con días en los que, en promedio, hay menos llamadas.

Estadísticos por tipo de día							
tipo_día	observaciones	llamadas_media	llamadas_mediana	llamadas_sd	total_llamadas	partidos_media	derrotas_media
Entre semana	9.369	6,709787597	6	4,484267885	62.864	0,04589604	0,017611271
Fin de semana	3.717	5,344901803	5	3,686220665	19.867	0,213075061	0,09039548



En cuanto a las variables futbolísticas, la base registra 1.159 días-provincia con actividad futbolística, 494 observaciones con derrota, 51 con derrota en descuento, 83 con remontada sufrida, 19 con derbi y 142 con partido de tensión alta. Estos valores muestran que algunas variables de alta intensidad emocional tienen un número de observaciones limitado, por lo que su interpretación debe realizarse con prudencia.

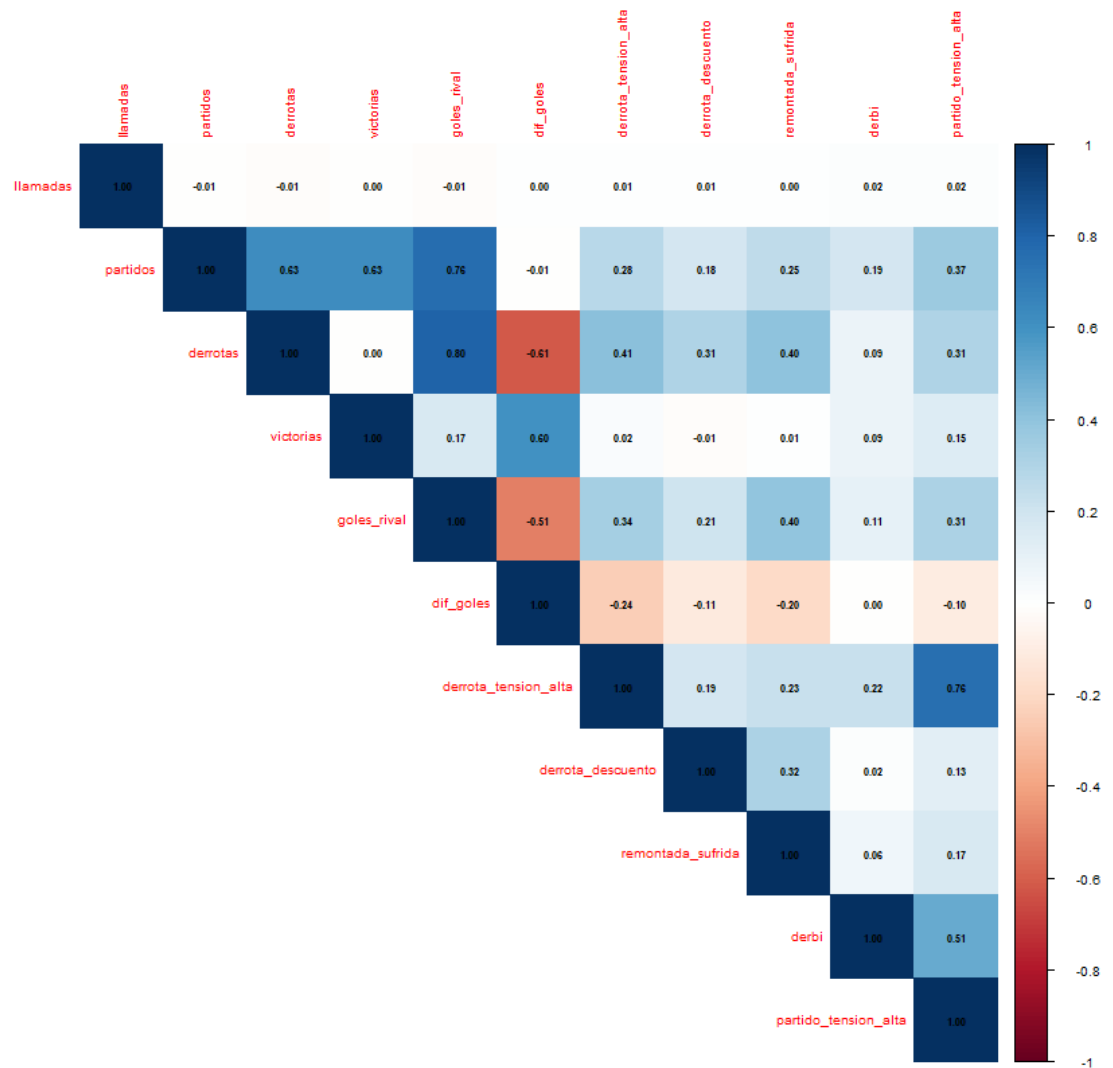
Descriptivos de variables futbolísticas						
partidos_media	derrotas_media	victorias_media	empates_media	goles_rival_media	goles_equipo_media	dif_goles_media
0,093382241	0,03828519	0,035534159	0,019562892	0,132508024	0,127999389	-0,004508635

Descriptivos de variables futbolísticas						
dias_con_futbol	dias_con_derrota	dias_con_derrota_descuento	dias_con_remontada	dias_con_derbi	dias_con_tension_alta	
1159	494		51	83	19	142

6.2.2. Matriz de correlaciones entre llamadas y variables futbolísticas

A continuación, se calcularon las correlaciones entre las llamadas y las principales variables futbolísticas. El objetivo de este análisis no es establecer causalidad, sino observar asociaciones lineales preliminares y detectar posibles relaciones entre las variables explicativas.

Correlaciones con llamadas		Correlaciones con llamadas	
variable	correlacion_con_llamadas	variable	correlacion_con_llamadas
llamadas	1	derrota_abultada	-0,005567034
partido_tension_alta	0,017587867	racha_derrotas_previa_max	-0,008568878
derbi	0,016113139	partido_alta_intensidad	-0,009496533
derrota_tension_alta	0,00766213	derrota_tardia	-0,009554123
derrota_descuento	0,007266075	partidos	-0,00989355
eliminacion	0,002991195	goles_equipo	-0,010255982
dif_goles	0,001934272	goles_rival	-0,011979628
remontada_sufrida	0,001397235	goles_totales	-0,012792182
empates	5,69E-04	derrotas	-0,013585161
victorias	-0,002494711		



Los resultados muestran que las correlaciones simples entre llamadas y variables futbolísticas son muy bajas. La correlación más elevada con llamadas corresponde a partido_tension_alta, con un valor de 0,018, seguida de derbi, con 0,016. Variables como derrota en descuento, derrota de tensión alta o eliminación presentan correlaciones positivas, pero prácticamente nulas. Por el contrario, variables generales como partidos, goles del rival, goles totales o derrotas presentan correlaciones ligeramente negativas, aunque también muy próximas a cero.

Este resultado es importante porque sugiere que no existe una relación lineal simple y fuerte entre fútbol y llamadas al 016. Sin embargo, esto no implica que no puedan existir asociaciones más específicas una vez se controlan factores temporales y territoriales. Las llamadas dependen de forma clara de la provincia y del tipo de día, por lo que las correlaciones simples pueden ocultar relaciones más complejas.

La matriz también muestra una elevada relación entre algunas variables futbolísticas. Por ejemplo, las derrotas están muy asociadas con los goles del rival, y los partidos de tensión alta se relacionan con los derbis y con las derrotas de tensión alta. Esta situación es esperable, ya que varias variables deportivas captan dimensiones relacionadas de un mismo evento. Por ello, en los modelos posteriores se evita introducir todas las variables simultáneamente y se trabaja con bloques de variables.

6.2.3. Comparación de medias entre días con y sin eventos deportivos

Como paso previo a la estimación econométrica, se realizaron comparaciones de medias entre días con y sin determinados eventos deportivos. Estas pruebas permiten comprobar si las diferencias observadas en Power BI aparecen también en los datos preparados para RStudio. (*Anexo: Tabla Comparación de medias eventos deportivos y Tabla Test de diferencia de medias*)

Los días con actividad futbolística registrada presentan una media de llamadas ligeramente superior a los días sin actividad. Entre semana, la diferencia es muy reducida: 6,77 llamadas en días con fútbol frente a 6,71 en días sin fútbol. En fin de semana, la diferencia es mayor: 5,69 llamadas en días con fútbol frente a 5,26 en días sin fútbol. El test de diferencia de medias resulta significativo en el total de la muestra y en fines de semana, pero no entre semana.

En cambio, las derrotas consideradas de forma general no muestran una asociación positiva con las llamadas. Entre semana, los días con derrota presentan una media de 6,58 llamadas, frente a 6,90 en días sin derrota. En fin de semana, ambas medias son prácticamente idénticas: 5,68 en días con derrota y 5,69 en días sin derrota. Los test de diferencia de medias no son significativos.

Las victorias tampoco presentan un patrón homogéneo. Entre semana, los días con victoria muestran una media algo menor que los días sin victoria. En fin de semana ocurre lo contrario: los días con victoria presentan 5,98 llamadas medias, frente a 5,51 en días sin victoria. Esta diferencia en fin de semana se aproxima a la significación estadística, pero no permite extraer una conclusión firme.

Las variables más relacionadas con intensidad emocional muestran resultados más interesantes. Los días con partido de tensión alta presentan medias superiores tanto entre semana como en fin de semana. Entre semana, la media es de 7,52 llamadas frente a 6,65 en días sin partido de tensión alta. En fin de semana, la media es de 6,41 frente a 5,60. En el total de la muestra, la diferencia resulta significativa, y en fin de semana se aproxima a la significación estadística.

También destacan los derbis. Aunque el número de observaciones es reducido, los días con derbi presentan medias claramente superiores, especialmente en fin de semana: 8,15 llamadas frente a 5,64. En este caso, el test de diferencia de medias resulta significativo en el total de la muestra y en fines de semana. Sin embargo, el bajo número de observaciones obliga a interpretar el resultado con prudencia.

En conjunto, el análisis de medias sugiere que la relación entre fútbol y llamadas al 016 no parece depender simplemente de que haya una derrota, sino de determinados contextos deportivos más específicos, como partidos de tensión alta o derbis.

6.3. Modelos sobre la base diaria

Para evaluar de forma más sistemática la asociación entre actividad futbolística y llamadas al 016, se estimaron distintos modelos de regresión en RStudio. La unidad principal de análisis es la provincia-fecha, y la variable dependiente es el número de llamadas pertinentes al 016

registradas en cada provincia y día.

La especificación general de los modelos lineales es la siguiente:

$$Llamadas_{it} = \beta_0 + \beta_1 \text{EventoFutbolístico}_{it} + \beta_2 \text{FinSemana}_t + \delta_i + \gamma_t + \varepsilon_{it}$$

donde i representa la provincia y t representa la fecha. Las variables futbolísticas recogen la actividad o el tipo de evento deportivo registrado en esa provincia y día. La variable fin_semana controla la diferencia entre días laborables y fines de semana. Los efectos de provincia (δ_i) controlan diferencias estructurales entre territorios, y los controles mes-año (γ_t) recogen variaciones temporales comunes.

Para reforzar la fiabilidad de los resultados, los modelos lineales se estimaron incorporando errores robustos HC1, lo que permite corregir posibles problemas de heterocedasticidad. Además, dado que la variable dependiente recoge un número de llamadas, es decir, valores enteros y no negativos, el análisis se complementó con otros enfoques de estimación de conteo como son los modelos de Poisson y binomial negativa.

La estrategia de estimación se organizó en bloques. En primer lugar, se estimó un modelo con la actividad futbolística general. En segundo lugar, se analizaron resultados deportivos básicos, como derrota, victoria, goles del rival y diferencia de goles. Después, se estimaron modelos centrados en variables de intensidad emocional, como derrotas en descuento, remontadas sufridas, derbis y partidos de tensión alta. Finalmente, se incluyeron variables retardadas para comprobar si la actividad futbolística del día anterior se asociaba con las llamadas del día siguiente.

6.4 Resultados de los modelos diarios

6.4.1 Modelos lineales con errores robustos

El primer modelo analiza la relación entre actividad futbolística registrada y llamadas al 016. Aunque en la comparación descriptiva los días con fútbol presentaban una media ligeramente superior de llamadas, esta relación desaparece al introducir controles por provincia, mes-año y fin de semana. El coeficiente de hay_futbol es 0,024 y no resulta estadísticamente significativo. En cambio, la variable fin_semana presenta un coeficiente negativo y altamente significativo, de -1,39. Esto confirma que el calendario semanal es un factor explicativo importante: los fines de semana se asocian con un menor número de llamadas, incluso después de controlar por provincia y periodo temporal.

El segundo modelo incorpora variables de resultado deportivo general. La variable hay_derrota presenta un coeficiente negativo de -0,59 y resulta estadísticamente significativa. Este resultado no apoya la hipótesis de que las derrotas, consideradas de forma general, aumenten las llamadas al 016. La variable hay_victoria no resulta significativa, ni tampoco los goles del rival. La diferencia de goles presenta un coeficiente negativo de -0,17 y se aproxima a la significación estadística, aunque debe interpretarse con cautela por su relación con otras variables del resultado deportivo.

Modelos lineales con errores robustos					
modelo	variable	estimate	std_error_robust	statistic	p_value
Modelo 1 - Actividad futbolística	hay_futbol	0,023676684	0,09904818	0,239042092	0,811076708
Modelo 1 - Actividad futbolística	fin_semana	-1,385479569	0,059806003	-23,16622925	2,22E-116
Modelo 2 - Resultado general	hay_derrota	-0,589311256	0,276894934	-2,128284715	0,033535479
Modelo 2 - Resultado general	hay_victoria	-0,071160251	0,291502207	-0,244115653	0,807186814
Modelo 2 - Resultado general	goles_rival	-0,013089155	0,11823469	-0,110704863	0,911870602
Modelo 2 - Resultado general	dif_goles	-0,172222797	0,096367426	-1,787147423	0,074189178
Modelo 2 - Resultado general	fin_semana	-1,207060138	0,199172329	-6,060380702	1,86E-09
Modelo 3 - Intensidad emocional	hay_derrota_descuento	-0,025078926	0,467259696	-0,05367235	0,957205967
Modelo 3 - Intensidad emocional	hay_remontada_sufrida	-0,075337403	0,362110026	-0,208051137	0,835227588
Modelo 3 - Intensidad emocional	hay_derbi	0,188335165	0,731864057	0,257336268	0,79696737
Modelo 3 - Intensidad emocional	hay_partido_tension_alta	0,757038954	0,414125289	1,828043287	0,067814035
Modelo 3 - Intensidad emocional	fin_semana	-1,149337352	0,19973651	-5,754267723	1,13E-08
Modelo 4 - Eventos negativos	hay_derrota_tension_alta	0,244102011	0,395340593	0,617447373	0,537067591
Modelo 4 - Eventos negativos	hay_derrota_descuento	0,004874817	0,473538742	0,010294442	0,991788237
Modelo 4 - Eventos negativos	hay_remontada_sufrida	-0,032727521	0,364795871	-0,089714615	0,928530348
Modelo 4 - Eventos negativos	hay_derrota_abultada	0,188218626	0,331762909	0,567328718	0,570606733
Modelo 4 - Eventos negativos	fin_semana	-1,178462473	0,19958424	-5,904586813	4,71E-09
Modelo 5 - Entre semana	hay_futbol	-0,096040207	0,207949173	-0,461844622	0,644203558
Modelo 5 - Entre semana	hay_derrota	-0,031137678	0,344963746	-0,090263624	0,928079671
Modelo 5 - Entre semana	hay_derrota_descuento	-0,388342954	0,807473315	-0,48093596	0,630573334
Modelo 5 - Entre semana	hay_remontada_sufrida	0,377514932	0,690862702	0,54643988	0,584776719
Modelo 5 - Entre semana	hay_partido_tension_alta	0,730191753	0,55608618	1,313090991	0,189184704
Modelo 6 - Fin de semana	hay_futbol	0,151177926	0,158182255	0,955719879	0,339277191
Modelo 6 - Fin de semana	hay_derrota	-0,201737425	0,228616269	-0,882428123	0,377603555
Modelo 6 - Fin de semana	hay_derrota_descuento	0,240647725	0,540301348	0,445395381	0,656060514
Modelo 6 - Fin de semana	hay_remontada_sufrida	0,058065763	0,417674652	0,139021516	0,889440831
Modelo 6 - Fin de semana	hay_partido_tension_alta	0,481088232	0,366810362	1,311544824	0,18975641
Modelo 7 - Retardos	hay_futbol	0,063274915	0,100460757	0,629847079	0,5288057
Modelo 7 - Retardos	hay_futbol_lag1	0,440947173	0,135667752	3,250198867	0,001156174
Modelo 7 - Retardos	hay_derrota_lag1	-0,26966633	0,209978943	-1,284254154	0,199075909
Modelo 7 - Retardos	hay_derrota_descuento_lag1	-0,51108038	0,55180861	-0,926191384	0,354363708
Modelo 7 - Retardos	hay_remontada_sufrida_lag1	0,653366977	0,427316778	1,528999117	0,126288973
Modelo 7 - Retardos	fin_semana	-1,406338969	0,060207086	-23,35836311	3,02E-118

En conjunto, el modelo de resultado general indica que las variables deportivas más básicas no captan de forma adecuada la posible relación entre fútbol y llamadas. Ni la existencia de una derrota ni la victoria ni el número de goles encajados muestran un patrón positivo claro.

El tercer modelo se centra en variables de intensidad emocional. En esta especificación se incluyen derrota en descuento, remontada sufrida, derbi y partido de tensión alta. El resultado más relevante es la variable hay_partido_tension_alta, que presenta un coeficiente positivo de 0,76 y una significación cercana a significativa de 0,07. Esto sugiere que los días con partidos de tensión alta podrían asociarse con un mayor número de llamadas, aunque la evidencia no alcanza el nivel convencional del 5%.

Las variables de derrota en descuento, remontada sufrida y derbi no resultan significativas en este modelo. Esto no implica que carezcan de interés descriptivo, pero sí indica que, al introducir controles territoriales y temporales, no se obtiene evidencia estadística suficiente para afirmar una asociación independiente.

El cuarto modelo analiza eventos negativos concretos, como derrota en partido de tensión alta, derrota en descuento, remontada sufrida y derrota abultada. Ninguna de estas variables resulta significativa. Por tanto, al aislar derrotas especialmente negativas, no se observa una asociación estadísticamente robusta con las llamadas.

También se estimaron modelos separados para entre semana y fin de semana. En estas especificaciones, ninguna de las variables deportivas principales resulta significativa. Esto indica que, aunque algunas diferencias descriptivas aparecen al separar por tipo de día, la evidencia se debilita al introducir controles adicionales.

Por último, el modelo con retardos ofrece un resultado relevante. La variable `hay_futbol_lag1`, que recoge si hubo actividad futbolística registrada el día anterior, presenta un coeficiente positivo de 0,44 y resulta estadísticamente significativa. Esto indica que la actividad futbolística del día anterior se asocia con un aumento de aproximadamente 0,44 llamadas al día siguiente, manteniendo constantes los controles del modelo.

Este resultado debe interpretarse con prudencia, pero es compatible con la posibilidad de que parte de la asociación entre fútbol y llamadas no se produzca necesariamente el mismo día del partido, sino en la jornada posterior. En cambio, los retardos de derrota, derrota en descuento y remontada sufrida no resultan significativos.

6.4.2 Modelos de conteo: Poisson y binomial negativa

Dado que la variable dependiente es un recuento de llamadas, se estimaron también modelos de conteo. En primer lugar se utilizaron modelos Poisson, y posteriormente modelos de binomial negativa para comprobar la fiabilidad de los resultados ante posible sobredispersión.

Sobredispersión de modelos Poisson	
modelo	dispersión
Poisson 1 - Actividad futbolística	1,432549297
Poisson 2 - Resultado general	1,321192932
Poisson 3 - Intensidad emocional	1,321954838

Los modelos Poisson presentan valores de sobredispersión entre 1,32 y 1,43. Estos valores indican que existe cierta sobredispersión, aunque no extremadamente elevada. Por este motivo, la binomial negativa se utiliza como especificación de contraste.

En el modelo Poisson de actividad futbolística general, `hay_futbol` no resulta significativo. En la binomial negativa tampoco resulta significativo, con un IRR de 1,017 y un p-valor de 0,301. Esto confirma que la actividad futbolística general no muestra una asociación robusta con las llamadas una vez incorporados los controles.

En el modelo de resultado general, la derrota vuelve a aparecer con signo negativo. En Poisson, el IRR de `hay_derrota` es 0,914 y resulta significativo. En binomial negativa, el IRR es 0,920 y se aproxima a la significación estadística. La diferencia de goles también presenta una asociación negativa y marginal en la binomial negativa. Estos resultados refuerzan la conclusión anterior: la derrota general no se asocia con un aumento de llamadas.

En el modelo de intensidad emocional, la variable más relevante vuelve a ser `hay_partido_tension_alta`. En Poisson, el IRR es 1,118 y resulta significativo (p-valor de 0,03). En binomial negativa, el IRR es 1,112 y se aproxima a la significación estadística (p-valor de 0,06). Esto implica que los días con partidos de tensión alta presentan un número esperado de llamadas aproximadamente un 11% superior, aunque la evidencia debe interpretarse como indicio y no como resultado definitivo.

El resto de variables de intensidad emocional, como derrota en descuento, remontada sufrida y derbi, no resultan significativas en los modelos de conteo. Por tanto, la conclusión es similar a la obtenida con los modelos lineales: los resultados no apoyan una relación general entre derrotas y llamadas, pero sí sugieren una posible asociación más débil y específica con los partidos de tensión alta.

Modelos de conteo - Variables clave					
term	estimate	std.error	statistic	p.value	modelo
hay_futbol	1,01920504	0,013344903	1,425484429	0,154017218	Poisson 1 - Actividad futbolística
fin_semana	0,791988999	0,00843232	-27,65642011	2,34E-168	Poisson 1 - Actividad futbolística
hay_derrota	0,913756493	0,038569276	-2,338419877	0,019365478	Poisson 2 - Resultado general
hay_victoria	1,002973154	0,038861589	0,076392731	0,939106653	Poisson 2 - Resultado general
goles_rival	0,998158282	0,015096576	-0,12210819	0,902813335	Poisson 2 - Resultado general
dif_goles	0,970412986	0,014219234	-2,112177071	0,034671264	Poisson 2 - Resultado general
fin_semana	0,820399024	0,025561422	-7,744656791	9,58E-15	Poisson 2 - Resultado general
hay_derrota_descuento	0,973300888	0,058834476	-0,459968517	0,645538818	Poisson 3 - Intensidad emocional
hay_remontada_sufrida	0,996025849	0,048964535	-0,081325584	0,935183028	Poisson 3 - Intensidad emocional
hay_derbi	1,040848263	0,094577525	0,423314298	0,672065964	Poisson 3 - Intensidad emocional
hay_partido_tension_alta	1,117833161	0,049526858	2,249125807	0,024504493	Poisson 3 - Intensidad emocional
fin_semana	0,827458594	0,025761539	-7,351898059	1,95E-13	Poisson 3 - Intensidad emocional
hay_futbol	1,016796157	0,016088489	1,035315461	0,300521612	Binomial negativa 1 - Actividad futbolística
fin_semana	0,794413741	0,01007991	-22,83263059	2,17E-115	Binomial negativa 1 - Actividad futbolística
hay_derrota	0,919675504	0,044206656	-1,89415784	0,058204062	Binomial negativa 2 - Resultado general
hay_victoria	1,002526035	0,044646663	0,056507025	0,9549379	Binomial negativa 2 - Resultado general
goles_rival	0,997673109	0,017339803	-0,134349977	0,893125835	Binomial negativa 2 - Resultado general
dif_goles	0,971264879	0,016188639	-1,801019693	0,071699776	Binomial negativa 2 - Resultado general
fin_semana	0,823009881	0,029401238	-6,625131585	3,47E-11	Binomial negativa 2 - Resultado general
hay_derrota_descuento	0,97720032	0,06819634	-0,338194279	0,735216787	Binomial negativa 3 - Intensidad emocional
hay_remontada_sufrida	0,997030965	0,056259341	-0,05285258	0,957849367	Binomial negativa 3 - Intensidad emocional
hay_derbi	1,040640955	0,111128562	0,358475138	0,719987771	Binomial negativa 3 - Intensidad emocional
hay_partido_tension_alta	1,112057533	0,057195015	1,857013808	0,06330919	Binomial negativa 3 - Intensidad emocional
fin_semana	0,829886862	0,029626396	-6,293910837	3,10E-10	Binomial negativa 3 - Intensidad emocional

6.4.3 Multicolinealidad y fiabilidad

Antes de interpretar los modelos como resultados finales, se revisó la multicolinealidad mediante indicadores VIF. Esta comprobación es importante porque muchas variables futbolísticas están relacionadas entre sí. Por ejemplo, una derrota suele estar asociada con goles en contra y diferencia de goles negativa.

Los resultados no muestran problemas severos de multicolinealidad. En el modelo de resultado general, el mayor valor ajustado corresponde a la diferencia de goles, con un GVIF ajustado de 2,17. El resto de variables se sitúa en valores inferiores o cercanos a 1,6. En los modelos de intensidad emocional, los valores también son moderados. Por tanto, aunque existe relación entre algunas variables deportivas, no se detectan niveles que invaliden la estimación.

Indicadores VIF de los Modelos				
modelo	variable	GVI	Df	GVI..1..2.Df..
Modelo 2 - Resultado general	hay_derrota	2,580007635	1	1,606240217
Modelo 2 - Resultado general	hay_victoria	2,598078703	1	1,611855671
Modelo 2 - Resultado general	goles_rival	2,493162903	1	1,57897527
Modelo 2 - Resultado general	dif_goles	4,692526751	1	2,166224077
Modelo 2 - Resultado general	fin_semana	1,097567293	1	1,047648459
Modelo 2 - Resultado general	factor(provincia)	1,40530647	6	1,028760436
Modelo 2 - Resultado general	factor(mes_ano)	1,659326458	48	1,005289061
Modelo 3 - Intensidad emocional	hay_derrota_descuento	1,133929246	1	1,064861139
Modelo 3 - Intensidad emocional	hay_remontada_sufrida	1,162618353	1	1,078247816
Modelo 3 - Intensidad emocional	hay_derbi	1,297577807	1	1,139112728
Modelo 3 - Intensidad emocional	hay_partido_tension_alta	1,986209702	1	1,409329522
Modelo 3 - Intensidad emocional	fin_semana	1,111118993	1	1,054096292
Modelo 3 - Intensidad emocional	factor(provincia)	1,320209013	6	1,023419194
Modelo 3 - Intensidad emocional	factor(mes_ano)	2,650561552	48	1,010205596
Modelo 4 - Eventos negativos	hay_derrota_tension_alta	1,288912582	1	1,135302859
Modelo 4 - Eventos negativos	hay_derrota_descuento	1,149560119	1	1,072175414
Modelo 4 - Eventos negativos	hay_remontada_sufrida	1,181437935	1	1,086939711
Modelo 4 - Eventos negativos	hay_derrota_abultada	1,09058577	1	1,044311146
Modelo 4 - Eventos negativos	fin_semana	1,10364506	1	1,050545125
Modelo 4 - Eventos negativos	factor(provincia)	1,321980409	6	1,023533555
Modelo 4 - Eventos negativos	factor(mes_ano)	1,89809866	48	1,00669788

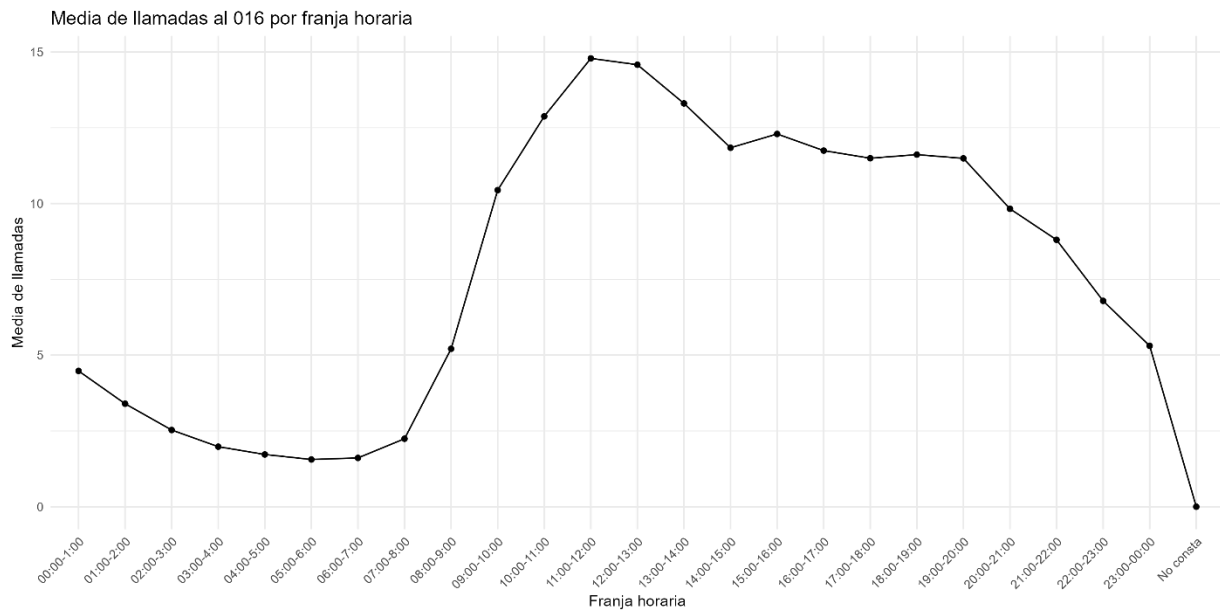
En términos de robustez, la evidencia más consistente es la ausencia de efecto de la actividad futbolística general y de las derrotas generales. En cambio, la variable de partido de tensión alta mantiene un signo positivo tanto en modelos lineales como en modelos de conteo, aunque con una significación estadística limitada. El resultado más destacado aparece en el modelo con retardos, donde la actividad futbolística registrada el día anterior se asocia de forma positiva y significativa con las llamadas del día siguiente. No obstante, este resultado debe interpretarse con cautela, ya que puede estar parcialmente relacionado con la propia dinámica semanal de las llamadas, especialmente con el mayor volumen observado al inicio de la semana.

6.5 Análisis exploratorio por franjas horarias

6.5.1 Estadísticos descriptivos de la base por franjas

El análisis por franjas horarias se plantea como un complemento exploratorio. La base contiene 9.841 observaciones, con una media de 8,41 llamadas por provincia, mes y franja horaria. En total se registran 3.035 partidos durante la franja, 1.237 partidos asignados a la franja posterior y 1.760 goles en contra registrados en franjas horarias.

Descriptivos generales de Franjas				
observaciones	llamadas_media	llamadas_mediana	llamadas_sd	llamadas_min
9841	8,406767605	6	7,281050059	0
llamadas_max	partidos_durante_tota	partidos_posterior_tota	goles_contra_franja_total	
39	3035	1237	1760	



Como vimos en el análisis descriptivo, la distribución horaria de llamadas muestra un patrón claro. Las llamadas son reducidas durante la madrugada y primeras horas de la mañana, aumentan a partir de las 08:00-09:00 y alcanzan sus máximos entre las 11:00 y las 13:00. En concreto, las franjas de 11:00-12:00 y 12:00-13:00 presentan medias de 14,79 y 14,58 llamadas, respectivamente. A partir de la tarde, el volumen se mantiene en niveles intermedios y desciende progresivamente durante la noche.

En cambio, la actividad futbolística se concentra principalmente en horario de tarde y noche. Los partidos durante la franja alcanzan sus valores más altos entre las 18:00 y las 23:00, especialmente en las franjas de 20:00-21:00, 21:00-22:00 y 22:00-23:00. Los goles en contra también se concentran en esas franjas, con un máximo en 22:00-23:00.

Descriptivos por Franja Horaria						
franja	observaciones	llamadas_media	llamadas_total	partidos_durante_tota	partidos_posterior_tota	goles_contra_franja_total
00:00-1:00	421	4,479809976	1886	1	31	0
01:00-2:00	402	3,39800995	1366	0	0	0
02:00-3:00	377	2,530503979	954	0	0	0
03:00-4:00	319	1,981191223	632	0	0	0
04:00-5:00	301	1,724252492	519	0	0	0
05:00-6:00	256	1,55859375	399	0	0	0
06:00-7:00	272	1,610294118	438	0	0	0
07:00-8:00	350	2,242857143	785	0	0	0
08:00-9:00	419	5,210023866	2183	0	0	0
09:00-10:00	440	10,44318182	4595	0	0	0
10:00-11:00	441	12,87755102	5679	0	0	0
11:00-12:00	441	14,78911565	6522	0	0	0
12:00-13:00	441	14,58049887	6430	7	0	4
13:00-14:00	441	13,30385488	5867	7	0	2
14:00-15:00	441	11,84126984	5222	117	7	75
15:00-16:00	439	12,29612756	5398	119	0	109
16:00-17:00	439	11,74715262	5157	197	117	114
17:00-18:00	441	11,49659864	5070	214	3	154
18:00-19:00	441	11,61451247	5122	438	194	117
19:00-20:00	441	11,49206349	5068	394	17	276
20:00-21:00	440	9,827272727	4324	466	259	211
21:00-22:00	439	8,804100228	3865	525	131	282
22:00-23:00	434	6,790322581	2947	477	79	361
23:00-00:00	434	5,306451613	2303	73	399	55
No consta	131	0	0	0	0	0

Esta diferencia horaria es relevante: las llamadas alcanzan sus máximos en horario diurno,

mientras que los partidos y goles se concentran en tarde-noche. Por tanto, desde una perspectiva puramente horaria, no se observa una coincidencia clara entre los momentos de mayor actividad futbolística y los momentos de mayor volumen de llamadas.

6.5.2 Correlaciones y modelos por franjas

Matriz de correlaciones por franjas									
	llamadas	partidos_durante	partidos_posterior	goles_totales_franja	goles_contra_franja	goles_favor_franja	derrotas_durante	derrotas_posterior	
llamadas	1	0,1221	0,0268	0,0846	0,0741	0,074	0,0879	0,018	
partidos_durante	0,1221	1	0,386	0,8135	0,7059	0,7188	0,7566	0,2792	
partidos_posterior	0,0268	0,386	1	0,2137	0,1715	0,2023	0,2868	0,7504	
goles_totales_franja	0,0846	0,8135	0,2137	1	0,8726	0,8789	0,6297	0,1443	
goles_contra_franja	0,0741	0,7059	0,1715	0,8726	1	0,5339	0,7212	0,1557	
goles_favor_franja	0,074	0,7188	0,2023	0,8789	0,5339	1	0,3857	0,0978	
derrotas_durante	0,0879	0,7566	0,2868	0,6297	0,7212	0,3857	1	0,3431	
derrotas_posterior	0,018	0,2792	0,7504	0,1443	0,1557	0,0978	0,3431	1	

Las correlaciones por franja son bajas. La correlación entre llamadas y partidos durante la franja es de 0,122. La correlación con goles totales en franja es de 0,085, con goles en contra es de 0,074 y con partidos en franja posterior es de 0,027. Estos valores indican que la asociación lineal entre llamadas y variables futbolísticas horarias es débil.

También se estimaron modelos exploratorios con la base por franjas. En estos modelos, los partidos durante la franja muestran una asociación positiva y significativa con las llamadas, mientras que los partidos en franja posterior aparecen con signo negativo. Sin embargo, estos resultados deben interpretarse con especial cautela, porque la base de llamadas por franja está agregada a nivel mensual y no permite vincular llamadas concretas con partidos concretos.

Modelos exploratorios por franjas					
modelo	variable	estimate	std_error_robust	statistic	p_value
Franjas 1 - Variables continuas	partidos_durante	1,200587874	0,093460416	12,84595052	1,82E-37
Franjas 1 - Variables continuas	partidos_posterior	-0,418994246	0,094897181	-4,415244376	1,02E-05
Franjas 1 - Variables continuas	goles_contra_franja	-0,190144552	0,104119072	-1,826222105	0,067847319
Franjas 2 - Variables binarias	hay_partidos_durante	3,222559687	0,196007281	16,44102031	6,18E-60
Franjas 2 - Variables binarias	hay_partidos_posterior	-0,558022294	0,180464036	-3,092152366	0,001992707
Franjas 2 - Variables binarias	hay_goles_contra_franja	-0,64737143	0,219181005	-2,953592761	0,003148466

En consecuencia, el análisis horario no se utiliza como evidencia principal del trabajo. Su principal aportación es mostrar que los patrones horarios de las llamadas y de la actividad futbolística son distintos. Las llamadas se concentran en horario diurno, mientras que los partidos, las franjas posteriores y los goles en contra se concentran en tarde-noche.

6.5.3 Ejercicio predictivo exploratorio

De forma complementaria, se exploró la posibilidad de construir modelos predictivos para identificar observaciones con llamadas especialmente elevadas. Para ello se utilizó una variable binaria de llamadas altas y se estimaron modelos de clasificación, concretamente una regresión logística y un Random Forest.

El modelo logit obtuvo un AUC de 0,597, mientras que el modelo Random Forest obtuvo un AUC de 0,494. Estos valores indican una capacidad predictiva limitada. El modelo logit apenas supera el umbral de azar, mientras que el Random Forest se sitúa prácticamente en el nivel de una clasificación aleatoria.

6.6 Conclusión del análisis estadístico y econométrico

El análisis estadístico permite matizar los resultados obtenidos en el análisis descriptivo. En primer lugar, los días con actividad futbolística registrada presentan una media de llamadas algo superior en el análisis simple, especialmente en fines de semana. Sin embargo, esta relación desaparece al introducir controles por provincia, mes-año y tipo de día. Por tanto, no se encuentra evidencia robusta de que la actividad futbolística general se asocie, por sí sola, con un aumento de llamadas al 016.

En segundo lugar, las derrotas consideradas de forma general no muestran una asociación positiva con las llamadas. Tanto en los modelos lineales como en los modelos de conteo, la variable de derrota aparece con signo negativo o no significativo. Esto contradice una interpretación simple según la cual perder un partido generaría automáticamente un aumento en las llamadas.

En tercer lugar, las variables de intensidad emocional ofrecen una lectura más matizada. Los partidos de tensión alta presentan un coeficiente positivo en los modelos lineales y un IRR superior a 1 en los modelos de conteo. Sin embargo, la significación estadística es limitada, especialmente en la binomial negativa. Por tanto, este resultado debe interpretarse como un indicio de asociación, no como una prueba concluyente.

El resultado más claro aparece en el modelo con retardos. La actividad futbolística registrada el día anterior se asocia de forma positiva y significativa con las llamadas del día siguiente. Este hallazgo sugiere que, si existe una relación entre fútbol y llamadas al 016, esta podría no concentrarse necesariamente en el mismo día del partido, sino aparecer parcialmente en la jornada posterior. Sin embargo, hay que tener en cuenta que el resultado puede estar sesgado por el mayor volumen de partidos en domingo frente a otros días de la semana y la mayor cantidad de llamadas los lunes.

El análisis por franjas horarias muestra que las llamadas y la actividad futbolística siguen patrones temporales distintos. Las llamadas se concentran en horario diurno, mientras que los partidos y goles se concentran en tarde-noche. Las correlaciones por franja son bajas, por lo que este bloque se mantiene como análisis exploratorio.

Finalmente, los modelos predictivos no muestran una capacidad suficiente para anticipar observaciones de llamadas altas. El AUC del modelo logit es de 0,597 y el del Random Forest de 0,494, por lo que no se incorporan como resultado principal.

En conjunto, los resultados no permiten afirmar que el fútbol, de forma general, incremente las llamadas al 016. Tampoco se observa una relación positiva clara con las derrotas ordinarias. La evidencia apunta más bien a una posible asociación débil con contextos deportivos de mayor tensión y, especialmente, con la actividad futbolística del día anterior. Por tanto, la relación entre fútbol y llamadas al 016, si existe, parece depender más del contexto concreto del evento deportivo y de su posible efecto temporal retardado que del resultado deportivo simple.

7. Discusión y conclusiones

7.1 Síntesis y discusión de los resultados

El objetivo de este trabajo era analizar si determinados eventos futbolísticos podían asociarse con variaciones de corto plazo en las llamadas pertinentes al Servicio 016 en Andalucía y la Región de Murcia. Para ello, se construyó una base de datos propia que combina llamadas diarias y por franja horaria con información de partidos, goles y resultados deportivos.

Los resultados muestran que no existe una relación simple entre fútbol y llamadas al 016. En el análisis descriptivo, los días con actividad futbolística registrada presentan una media de llamadas ligeramente superior a los días sin fútbol, especialmente en fin de semana. Sin embargo, esta relación deja de ser significativa cuando se introducen controles por provincia, mes-año y tipo de día. Por tanto, la existencia de fútbol en general no parece explicar por sí sola un aumento claro de llamadas.

Tampoco se encuentra evidencia de que las derrotas ordinarias se asocien con un incremento de llamadas. De hecho, en los modelos estimados, la variable de derrota aparece con signo negativo o no significativo. Esto sugiere que el resultado deportivo por sí solo no capta adecuadamente la posible carga emocional del evento.

En cambio, algunas variables más específicas sí muestran patrones más interesantes. En el análisis descriptivo destacan los partidos de tensión alta, los derbis, las remontadas sufridas y las derrotas en descuento. En los modelos econométricos, la variable de partido de tensión alta mantiene un signo positivo, aunque con una significación estadística limitada. Por tanto, este resultado debe interpretarse como un indicio, no como una prueba concluyente.

El resultado más claro aparece en los modelos con retardos: la actividad futbolística registrada el día anterior se asocia de forma positiva y significativa con las llamadas del día siguiente. Esto sugiere que, si existe una relación entre fútbol y llamadas al 016, puede no producirse necesariamente el mismo día del partido, sino aparecer parcialmente en la jornada posterior. No obstante, este resultado puede estar influenciado por la dinámica de eventos futbolísticos a final de semana y de aumento de llamadas al comienzo de la misma, por lo que hay que interpretarlo con prudencia.

Por último, el análisis por franjas horarias muestra que las llamadas y la actividad futbolística siguen patrones temporales distintos. Las llamadas se concentran principalmente en horario diurno, mientras que los partidos y goles se concentran en la tarde y noche. Por ello, el análisis horario se mantiene como exploratorio.

7.2 Relación con la literatura previa

Los resultados del trabajo coinciden parcialmente con la literatura previa sobre eventos deportivos y violencia doméstica. Estudios internacionales han señalado que ciertos acontecimientos deportivos pueden actuar como desencadenantes de corto plazo, especialmente cuando generan frustración, tensión emocional o consumo de alcohol.

Sin embargo, este trabajo no encuentra una relación clara entre derrotas generales y llamadas al 016. La evidencia apunta más bien a que el contexto del partido puede ser más importante

que el resultado en sí. Es decir, no todas las derrotas parecen tener la misma relevancia: los partidos más intensos, simbólicos o decisivos podrían tener una mayor capacidad de asociarse con variaciones en las llamadas.

También debe tenerse en cuenta que este estudio utiliza llamadas al 016 como indicador indirecto. A diferencia de otros trabajos basados en denuncias o registros policiales, las llamadas reflejan demanda de información, asesoramiento o ayuda, pero no equivalen necesariamente a episodios confirmados de violencia. Por ello, los resultados no son directamente comparables, aunque sí aportan una primera aproximación al caso español.

7.3 Implicaciones, limitaciones y líneas futuras

Desde el punto de vista de las políticas públicas, los resultados sugieren que determinados eventos deportivos de alta tensión podrían considerarse momentos de especial atención preventiva. Esto no significa que el fútbol sea una causa directa de violencia de género, ni que todos los partidos representen un riesgo. Más bien, algunos encuentros especialmente intensos podrían formar parte de un conjunto más amplio de factores contextuales a tener en cuenta.

En este sentido, grandes derbis, eliminatorias o partidos decisivos podrían ser momentos adecuados para reforzar la visibilidad del 016 y otros recursos de ayuda. Esta recomendación debe plantearse con prudencia, evitando estigmatizar a los aficionados y entendiendo el fútbol como un posible contexto de tensión, no como causa directa.

El estudio presenta varias limitaciones: Las llamadas diarias solo están disponibles por provincia y fecha, y las llamadas por franja horaria están agregadas mensualmente. Esto impide analizar con precisión qué ocurre durante el partido o en las horas inmediatamente posteriores. Además, la base futbolística procede de eventos de gol, por lo que los partidos 0-0 no aparecen registrados. Además, el análisis se realiza a nivel provincial, aunque el efecto emocional de un equipo puede concentrarse en una ciudad o área concreta. Por último, existen factores no observados, como alcohol, clima, festivos, noticias o campañas institucionales, que también pueden influir en las llamadas.

Como líneas futuras, sería útil disponer de datos del 016 por provincia, día y hora, incorporar partidos sin goles, medir la importancia deportiva de cada encuentro y ampliar el análisis al conjunto de España.

7.4 Conclusión final

Este trabajo ha analizado la posible relación entre actividad futbolística y llamadas al Servicio 016 en Andalucía y la Región de Murcia. Los resultados no permiten afirmar que el fútbol, de forma general, incremente las llamadas, ni que las derrotas ordinarias estén asociadas a un aumento sistemático.

La evidencia apunta a una relación más matizada. Algunos contextos deportivos de mayor tensión, como derbis, eliminatorias o partidos decisivos, podrían estar asociados con mayores niveles de llamadas, aunque la evidencia estadística no es concluyente. Además, la actividad futbolística del día anterior, sí aparece relacionada con un aumento de llamadas al día siguiente, aunque se debe estudiar más en detalle, ya que este resultado está muy

influenciado por la concentración de las llamadas al 016 y eventos futbolísticos en el comienzo y el final de la semana, respectivamente.

En definitiva, el estudio aporta una primera aproximación al contexto español y muestra que la relación entre fútbol y llamadas al 016, si existe, no depende simplemente del resultado deportivo, sino del tipo de partido, su carga emocional y el momento temporal en que se analiza. Los resultados deben interpretarse con cautela, pero abren una línea de investigación útil para seguir estudiando cómo ciertos eventos sociales pueden relacionarse con la demanda de ayuda frente a la violencia de género.

8. Bibliografía

Card, D., & Dahl, G. B. (2011). Family violence and football: The effect of unexpected emotional cues on violent behavior. *The Quarterly Journal of Economics*, 126(1), 103–143. <https://doi.org/10.1093/qje/qjr001>

Centers for Disease Control and Prevention. (2024, February 8). *Risk and protective factors: Intimate partner violence prevention*. U.S. Department of Health and Human Services. <https://www.cdc.gov/intimate-partner-violence/risk-factors/index.html>

Delegación del Gobierno contra la Violencia de Género. (s. f.). *Teléfono 016 - 016 online - WhatsApp*. Ministerio de Igualdad. Recuperado el 1 de junio de 2026, de <https://violenciagenero.igualdad.gob.es/informacion-3/recursos/telefono016/>

Jefatura del Estado. (2004, 29 de diciembre). *Ley Orgánica 1/2004, de 28 de diciembre, de Medidas de Protección Integral contra la Violencia de Género*. Boletín Oficial del Estado, núm. 313. <https://www.boe.es/buscar/act.php?id=BOE-A-2004-21760>

Kirby, S., Francis, B., & O'Flaherty, R. (2014). Can the FIFA World Cup football (soccer) tournament be associated with an increase in domestic abuse? *Journal of Research in Crime and Delinquency*, 51(3), 259–276. <https://doi.org/10.1177/0022427813494843>

Observatorio Estatal de Violencia sobre la Mujer. (2022). *XV Informe anual del Observatorio Estatal de Violencia sobre la Mujer: Resumen ejecutivo*. Ministerio de Igualdad. https://violenciagenero.igualdad.gob.es/wp-content/uploads/RE_XV_2021-1.pdf

World Health Organization. (2019). *RESPECT women: Preventing violence against women* (WHO/RHR/18.19). <https://www.who.int/publications/i/item/WHO-RHR-18.19>

World Health Organization. (2024, March 25). *Violence against women*. <https://www.who.int/news-room/fact-sheets/detail/violence-against-women>

9. Anexos

9.1 Script de ejemplo de construcción de la base futbolística: Sevilla FC, temporada 20-21

```
import requests
import pandas as pd
import time

API_KEY = "INTRODUCIR_API_KEY"

headers = {
    "x-apisports-key": API_KEY
}

EQUIPO_REFERENCIA = "Sevilla FC" # se cambia
TEAM_ID = 536 # se cambia

SEASON = 2020 # se cambia
TEMPORADA = "2020-21" # se cambia

DIVISION = "1ª" # se cambia
CCAA = "Andalucía" # se cambia

url_fixtures = "https://v3.football.api-sports.io/fixtures"
url_events = "https://v3.football.api-sports.io/fixtures/events"

LEAGUES = {
    "La Liga": 140,
    "Copa del Rey": 143,
    "Supercopa": 556,
    "Champions League": 2,
    "Europa League": 3
}

def franja_30_minutos(fecha):
    hora = fecha.hour
    minuto = fecha.minute

    inicio_min = (minuto // 30) * 30
    fin_min = inicio_min + 30

    inicio = f"{hora:02d}:{inicio_min:02d}"

    if fin_min == 60:
        fin = f"{hora+1:02d}:00"
    else:
        fin = f"{hora:02d}:{fin_min:02d}"

    return f"{inicio}-{fin}"

def franja_gol(fecha, minuto_gol):
    total_min = fecha.hour * 60 + fecha.minute + minuto_gol

    hora = total_min // 60
    minuto = total_min % 60
```

```

inicio_min = (minuto // 30) * 30
fin_min = inicio_min + 30

inicio = f"{hora:02d}:{inicio_min:02d}"

if fin_min == 60:
    fin = f"{hora+1:02d}:00"
else:
    fin = f"{hora:02d}:{fin_min:02d}"

return f"{inicio}-{fin}"
meses_es = {
    1: "Enero",
    2: "Febrero",
    3: "Marzo",
    4: "Abril",
    5: "Mayo",
    6: "Junio",
    7: "Julio",
    8: "Agosto",
    9: "Septiembre",
    10: "Octubre",
    11: "Noviembre",
    12: "Diciembre"
}
df_total = pd.DataFrame()

for nombre_liga, league_id in LEAGUES.items():

    print(f"Descargando {nombre_liga}...")

    params_fixtures = {
        "team": TEAM_ID,
        "league": league_id,
        "season": SEASON
    }

    response = requests.get(
        url_fixtures,
        headers=headers,
        params=params_fixtures
    )

    fixtures_data = response.json()

    if fixtures_data["results"] == 0:
        continue

    filas = []

    for partido in fixtures_data["response"]:

```

```

fixture_id = partido["fixture"]["id"]
fecha_dt = pd.to_datetime(partido["fixture"]["date"]).tz_convert(None)

equipo_local = partido["teams"]["home"]["name"]
equipo_visitante = partido["teams"]["away"]["name"]
competicion = partido["league"]["name"]
jornada_raw = partido["league"]["round"]

if "Regular Season" in jornada_raw:
    jornada = "J" + jornada_raw.split("-")[-1].strip()
else:
    jornada = jornada_raw

goles_home = partido["goals"]["home"]
goles_away = partido["goals"]["away"]

resultado = f"{goles_home}-{goles_away}"

hora_inicio = fecha_dt.strftime("%H:%M")
franja_partido = franja_30_minutos(fecha_dt)

params_events = {
    "fixture": fixture_id
}

response_events = requests.get(
    url_events,
    headers=headers,
    params=params_events
)

events_data = response_events.json()

time.sleep(0.3)

for e in events_data["response"]:

    if e["type"] == "Goal":

        equipo_marcador = e["team"]["name"]
        minuto = e["time"]["elapsed"]
        extra = e["time"]["extra"]

        minuto_gol = f"{minuto}+{extra}" if extra else minuto

        local_visit_marcador = "Local" if equipo_marcador ==
equipo_local else "Visitante"
        favor_contra = "A FAVOR" if equipo_marcador ==
EQUIPO_REFERENCIA else "EN CONTRA"

        filas.append({
            "Temporada": TEMPORADA,
            "Equipo referencia": EQUIPO_REFERENCIA,

```

```

        "División": DIVISION,
        "Mes": meses_es[fecha_dt.month],
        "Nº Mes": fecha_dt.month,
        "Año": fecha_dt.year,
        "Fecha": fecha_dt.strftime("%d/%m/%Y"),
        "Jornada": jornada,
        "Equipo Local": equipo_local,
        "Equipo Visitante": equipo_visitante,
        "Equipo Marcador": equipo_marcador,
        "Resultado": resultado,
        "Minuto Gol": minuto_gol,
        "Franja Partido": franja_partido,
        "Hora Inicio": hora_inicio,
        "Franja Gol": franja_gol(fecha_dt, minuto),
        "Local/Visit.": local_visit_marcador,
        "Competición": competicion,
        "Notas": "",
        "CCAA": CCAA,
        "Gol A FAVOR / EN CONTRA": favor_contra
    })

```

```

df_liga = pd.DataFrame(filas)
df_total = pd.concat([df_total, df_liga], ignore_index=True)

```

```

df_total.head()
nombre_archivo = f"{EQUIPO_REFERENCIA.replace(' ',
'_' )}_{TEMPORADA}_COMPLETO.xlsx"

```

```

df_total.to_excel(
    nombre_archivo,
    index=False
)

```

```

df_total["Competición"].value_counts()
df_total.shape

```

9.2 Tablas auxiliares y ejemplos de datasets finales

9.2.1 Tablas auxiliares: Maestro Equipo-Provincia-CCAA, Maestro Franjas y Calendario

Maestro Equipo-Provincia-CCAA

Equipo referencia	CCAA	Provincia
Almería	Andalucía	Almería
Cádiz	Andalucía	Cádiz
Córdoba	Andalucía	Córdoba
Granada CF	Andalucía	Granada
Málaga CF	Andalucía	Málaga
Real Betis	Andalucía	Sevilla
Sevilla FC	Andalucía	Sevilla
FC Cartagena	Región de Murcia	Murcia

Maestro Franjas

Franja 30 min	Franja 1h
00:00-00:30	00:00-1:00
00:30-1:00	00:00-1:00
1:00-1:30	1:00-2:00
1:30-2:00	1:00-2:00
2:00-2:30	2:00-3:00
2:30-3:00	2:00-3:00
3:00-3:30	3:00-4:00
3:30-4:00	3:00-4:00
4:00-4:30	4:00-5:00
4:30-5:00	4:00-5:00
5:00-5:30	5:00-6:00
5:30-6:00	5:00-6:00
6:00-6:30	6:00-7:00
6:30-7:00	6:00-7:00
7:00-7:30	7:00-8:00
7:30-8:00	7:00-8:00
8:00-8:30	8:00-9:00
8:30-9:00	8:00-9:00
9:00-9:30	9:00-10:00
9:30-10:00	9:00-10:00
10:00-10:30	10:00-11:00
10:30-11:00	10:00-11:00
11:00-11:30	11:00-12:00
11:30-12:00	11:00-12:00
12:00-12:30	12:00-13:00
12:30-13:00	12:00-13:00
13:00-13:30	13:00-14:00
13:30-14:00	13:00-14:00
14:00-14:30	14:00-15:00
14:30-15:00	14:00-15:00
15:00-15:30	15:00-16:00
15:30-16:00	15:00-16:00
16:00-16:30	16:00-17:00
16:30-17:00	16:00-17:00
17:00-17:30	17:00-18:00
17:30-18:00	17:00-18:00
18:00-18:30	18:00-19:00
18:30-19:00	18:00-19:00
19:00-19:30	19:00-20:00
19:30-20:00	19:00-20:00
20:00-20:30	20:00-21:00
20:30-21:00	20:00-21:00
21:00-21:30	21:00-22:00
21:30-22:00	21:00-22:00
22:00-22:30	22:00-23:00
22:30-23:00	22:00-23:00
23:00-23:30	23:00-00:00
23:30-00:00	23:00-00:00

Date	key_date	mes	Mes mmm	Mes- Año	Tipo Dia	DiaSemana
01/01/2025 0:00:00	20250101	1	ENE	1-2025	Entre semana	miércoles
02/01/2025 0:00:00	20250102	1	ENE	1-2025	Entre semana	jueves
03/01/2025 0:00:00	20250103	1	ENE	1-2025	Entre semana	viernes
04/01/2025 0:00:00	20250104	1	ENE	1-2025	Fin de semana	sábado
05/01/2025 0:00:00	20250105	1	ENE	1-2025	Fin de semana	domingo
06/01/2025 0:00:00	20250106	1	ENE	1-2025	Entre semana	lunes
07/01/2025 0:00:00	20250107	1	ENE	1-2025	Entre semana	martes
08/01/2025 0:00:00	20250108	1	ENE	1-2025	Entre semana	miércoles
09/01/2025 0:00:00	20250109	1	ENE	1-2025	Entre semana	jueves
10/01/2025 0:00:00	20250110	1	ENE	1-2025	Entre semana	viernes
11/01/2025 0:00:00	20250111	1	ENE	1-2025	Fin de semana	sábado
12/01/2025 0:00:00	20250112	1	ENE	1-2025	Fin de semana	domingo
13/01/2025 0:00:00	20250113	1	ENE	1-2025	Entre semana	lunes
14/01/2025 0:00:00	20250114	1	ENE	1-2025	Entre semana	martes
15/01/2025 0:00:00	20250115	1	ENE	1-2025	Entre semana	miércoles
16/01/2025 0:00:00	20250116	1	ENE	1-2025	Entre semana	jueves
17/01/2025 0:00:00	20250117	1	ENE	1-2025	Entre semana	viernes
18/01/2025 0:00:00	20250118	1	ENE	1-2025	Fin de semana	sábado
19/01/2025 0:00:00	20250119	1	ENE	1-2025	Fin de semana	domingo
20/01/2025 0:00:00	20250120	1	ENE	1-2025	Entre semana	lunes
21/01/2025 0:00:00	20250121	1	ENE	1-2025	Entre semana	martes
22/01/2025 0:00:00	20250122	1	ENE	1-2025	Entre semana	miércoles
23/01/2025 0:00:00	20250123	1	ENE	1-2025	Entre semana	jueves
24/01/2025 0:00:00	20250124	1	ENE	1-2025	Entre semana	viernes
25/01/2025 0:00:00	20250125	1	ENE	1-2025	Fin de semana	sábado

Hay Derrota Tardía	Hay Derrota Descuento	Hay Remontada Sufrida	Hay Derrota Abultada	Hay Derbi	Hay Eliminación	Hay Partido Nocturno	Hay Alta Intensidad
Sin derrota tardía	Sin derrota descuento	Sin remontada sufrida	Sin derrota abultada	Sin derbi	Sin eliminación	Con partido nocturno	Sin alta intensidad
Sin derrota tardía	Sin derrota descuento	Sin remontada sufrida	Sin derrota abultada	Sin derbi	Sin eliminación	Sin partido nocturno	Sin alta intensidad
Sin derrota tardía	Sin derrota descuento	Sin remontada sufrida	Sin derrota abultada	Sin derbi	Sin eliminación	Sin partido nocturno	Sin alta intensidad
Sin derrota tardía	Sin derrota descuento	Sin remontada sufrida	Sin derrota abultada	Sin derbi	Sin eliminación	Sin partido nocturno	Sin alta intensidad
Sin derrota tardía	Sin derrota descuento	Sin remontada sufrida	Sin derrota abultada	Sin derbi	Sin eliminación	Sin partido nocturno	Sin alta intensidad
Sin derrota tardía	Sin derrota descuento	Sin remontada sufrida	Sin derrota abultada	Sin derbi	Sin eliminación	Sin partido nocturno	Sin alta intensidad
Sin derrota tardía	Sin derrota descuento	Sin remontada sufrida	Sin derrota abultada	Sin derbi	Sin eliminación	Sin partido nocturno	Sin alta intensidad

Grupo Derrotas	Grupo Partidos	Grupo Goles Contra	Grupo Racha Previa	DiaSem_Numero	Hay Partido Tension Alta
1 derrota	1 partido	2 goles en contra	2 derrotas previas		1 Sin partido tensión alta
0 derrotas	0 partidos	0 goles en contra	Sin racha		1 Sin partido tensión alta
0 derrotas	0 partidos	0 goles en contra	Sin racha		1 Sin partido tensión alta
0 derrotas	0 partidos	0 goles en contra	Sin racha		1 Sin partido tensión alta
0 derrotas	0 partidos	0 goles en contra	Sin racha		1 Sin partido tensión alta
0 derrotas	0 partidos	0 goles en contra	Sin racha		1 Sin partido tensión alta
0 derrotas	0 partidos	0 goles en contra	Sin racha		1 Sin partido tensión alta

Ejemplo Dataset_R_Franjas (abr/2024, 21:00-22:00)

Año	Mes	Mes_Numero	Mes-Año	Provincia	CCAA	Franja Cruce	Llamadas	Partidos Franja Posterior	Derrotas Franja Posterior
Año 2024	Abril	4	4-2024	Sevilla	Andalucía	21:00-22:00	19	0	0
Año 2024	Abril	4	4-2024	Granada	Andalucía	21:00-22:00	3	0	0
Año 2024	Abril	4	4-2024	Almería	Andalucía	21:00-22:00	2	0	0
Año 2024	Abril	4	4-2024	Murcia	Región de Murcia	21:00-22:00	24	0	0
Año 2024	Abril	4	4-2024	Cádiz	Andalucía	21:00-22:00	5	0	0
Año 2024	Abril	4	4-2024	Málaga	Andalucía	21:00-22:00	10	0	0
Año 2024	Abril	4	4-2024	Córdoba	Andalucía	21:00-22:00	8	0	0

Victorias Franja Posterior	Empates Franja Posterior	Derrota Tardía Franja Posterior	Derrota Descuento Franja Posterior	Derrota Abultada Franja Posterior	Remontada Sufrida Franja Posterior	Derbi Franja Posterior
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0

Eliminacion Franja Posterior	Partido Alta Intensidad Franja Posterior	Partido Tension Alta Franja Posterior	Derrota Tension Alta Franja Posterior	Goles Favor Franja Gol	Goles Contra Franja Gol	Goles Totales Franja Gol
0	0	0	0	1	1	2
0	0	0	0	1	2	3
0	0	0	0	1	1	2
0	0	0	0	2	2	4
0	0	0	0	0	3	3
0	0	0	0	0	0	0
0	0	0	0	0	0	0

Franja Inicio Min	Franja Fin Min Abs	Partidos Durante	Victorias Durante	Derrotas Durante	Empates Durante	Goles Equipo Durante	Goles Rival Durante
12:00:00 AM	12:00:00 AM	4	2	0	2	6	4
12:00:00 AM	12:00:00 AM	2	0	1	1	1	2
12:00:00 AM	12:00:00 AM	1	0	0	1	2	2
12:00:00 AM	12:00:00 AM	2	1	1	0	2	2
12:00:00 AM	12:00:00 AM	2	0	2	0	1	5
12:00:00 AM	12:00:00 AM		0	0	0	0	0
12:00:00 AM	12:00:00 AM		0	0	0	0	0

Dif Goles Durante	Goles Contra 80+ Durante	Goles Contra 90+ Durante	Goles Favor 80+ Durante	Derrota Tardía Durante	Derrota Descuento Durante	Derrota Abultada Durante	Remontada Sufrida Durante	Derbi Durante
2	2	2	1	0	0	0	0	2
-1	0	0	0	0	0	0	0	0
0	0	0	1	0	0	0	0	0
0	0	0	0	0	0	0	0	0
-4	1	0	1	0	0	0	1	0
0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0

Eliminacion Durante	Derrota Local Durante	Derrota Visitante Durante	Encaja Primero Durante	Derrota Encajando Primero Durante	Gol Contra Temprano Durante	Derrota Gol Contra Temprano Durante
0	0	0	1	0	0	0
0	1	0	1	1	0	0
0	0	0	0	0	0	0
0	1	0	1	1	0	0
0	1	1	2	2	1	1
0	0	0	0	0	0	0
0	0	0	0	0	0	0

Derrota Con Racha Previa Durante	Partido Alta Intensidad Durante	Partido Tension Alta Durante	Derrota Tension Alta Durante	Hay Futbol Durante	Hay Futbol Previo
0	0	2	0	Con actividad futbolística	Sin actividad futbolística
1	0	0	0	Con actividad futbolística	Sin actividad futbolística
0	0	0	0	Con actividad futbolística	Sin actividad futbolística
0	0	0	0	Con actividad futbolística	Sin actividad futbolística
0	1	0	0	Con actividad futbolística	Sin actividad futbolística
0	0	0	0	Sin actividad futbolística	Sin actividad futbolística
0	0	0	0	Sin actividad futbolística	Sin actividad futbolística

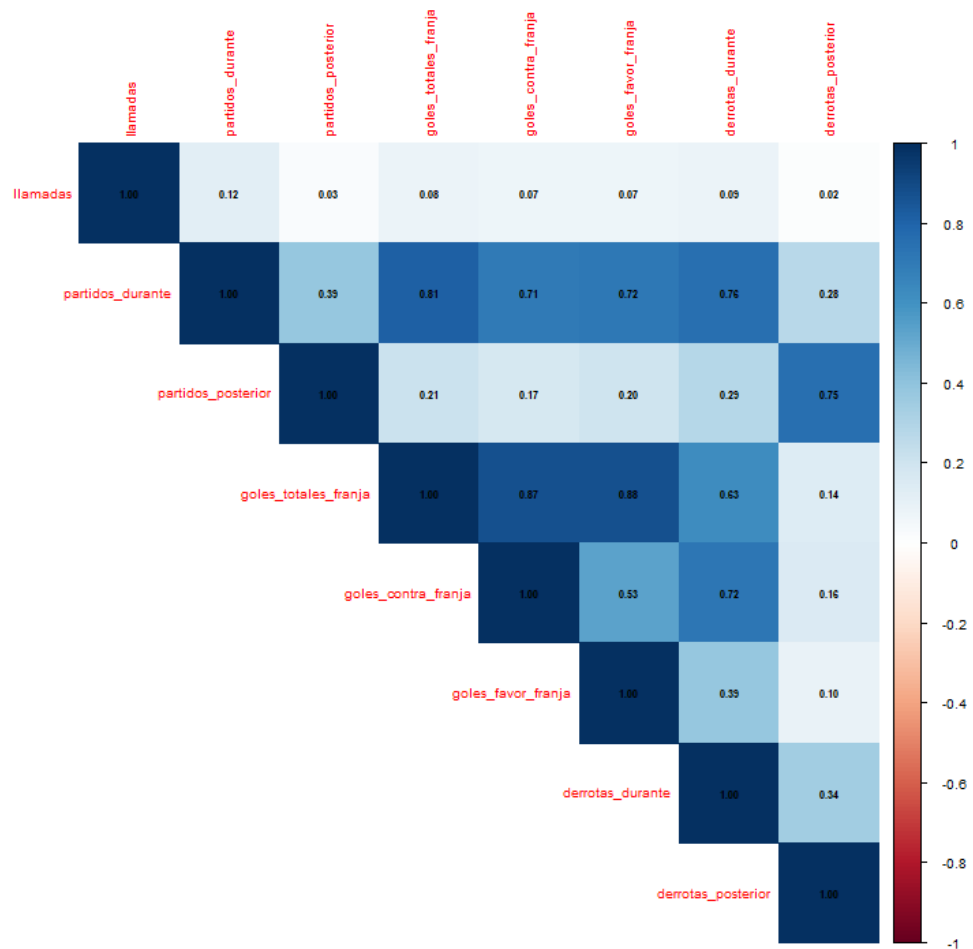
9.3 Gráficos adicionales del análisis estadístico y econométrico en Rstudio

Comparación de medias por eventos deportivos

tipo_dia	valor_evento	variable	n	llamadas_media	llamadas_sd
Entre semana	0	hay_futbol	8946	6,706908115	4,487850976
Entre semana	1	hay_futbol	423	6,770685579	4,412594132
Fin de semana	0	hay_futbol	2981	5,26031533	3,694817489
Fin de semana	1	hay_futbol	736	5,6875	3,633578285
Entre semana	0	hay_derrota	258	6,895348837	4,611006296
Entre semana	1	hay_derrota	165	6,575757576	4,089175949
Fin de semana	0	hay_derrota	407	5,692874693	3,686818932
Fin de semana	1	hay_derrota	329	5,680851064	3,572203563
Entre semana	0	hay_victoria	246	6,865853659	4,359635433
Entre semana	1	hay_victoria	177	6,638418079	4,494225308
Fin de semana	0	hay_victoria	460	5,510869565	3,539058641
Fin de semana	1	hay_victoria	276	5,981884058	3,774030602
Entre semana	0	hay_partido_tension_alta	363	6,64738292	4,247477319
Entre semana	1	hay_partido_tension_alta	60	7,516666667	5,283462099
Fin de semana	0	hay_partido_tension_alta	654	5,596330275	3,593304816
Fin de semana	1	hay_partido_tension_alta	82	6,414634146	3,887146338
Entre semana	0	hay_derrota_tension_alta	383	6,759791123	4,405464474
Entre semana	1	hay_derrota_tension_alta	40	6,875	4,535826049
Fin de semana	0	hay_derrota_tension_alta	691	5,643994211	3,635657263
Fin de semana	1	hay_derrota_tension_alta	45	6,355555556	3,574884331
Entre semana	0	hay_derrota_descuento	401	6,713216958	4,378361551
Entre semana	1	hay_derrota_descuento	22	7,818181818	4,991768115
Fin de semana	0	hay_derrota_descuento	707	5,6718529	3,639543187
Fin de semana	1	hay_derrota_descuento	29	6,068965517	3,524719183
Entre semana	0	hay_derrota_tardia	370	6,751351351	4,434749552
Entre semana	1	hay_derrota_tardia	53	6,905660377	4,293394528
Fin de semana	0	hay_derrota_tardia	638	5,728840125	3,676069103
Fin de semana	1	hay_derrota_tardia	98	5,418367347	3,348868002
Entre semana	0	hay_remontada_sufrida	401	6,74563591	4,41532979
Entre semana	1	hay_remontada_sufrida	22	7,227272727	4,439345921
Fin de semana	0	hay_remontada_sufrida	675	5,65037037	3,627599325
Fin de semana	1	hay_remontada_sufrida	61	6,098360656	3,704523892
Entre semana	0	hay_derrota_abultada	392	6,836734694	4,483430997
Entre semana	1	hay_derrota_abultada	31	5,935483871	3,336020422
Fin de semana	0	hay_derrota_abultada	677	5,652880355	3,618539312
Fin de semana	1	hay_derrota_abultada	59	6,084745763	3,811453409
Entre semana	0	hay_derbi	417	6,745803357	4,386675398
Entre semana	1	hay_derbi	6	8,5	6,220932406
Fin de semana	0	hay_derbi	723	5,643153527	3,636140082
Fin de semana	1	hay_derbi	13	8,153846154	2,544476169
Entre semana	0	hay Eliminacion	399	6,746867168	4,388608364
Entre semana	1	hay Eliminacion	24	7,166666667	4,878227286
Fin de semana	0	hay Eliminacion	725	5,693793103	3,647072176
Fin de semana	1	hay Eliminacion	11	5,272727273	2,686667419
Entre semana	0	hay_derrota_con_racha_previa	392	6,711734694	4,401355877
Entre semana	1	hay_derrota_con_racha_previa	31	7,516129032	4,55975853
Fin de semana	0	hay_derrota_con_racha_previa	687	5,729257642	3,687430767
Fin de semana	1	hay_derrota_con_racha_previa	49	5,102040816	2,732862495
Entre semana	0	hay_derrota_encajando_primero	278	6,913669065	4,60572868
Entre semana	1	hay_derrota_encajando_primero	145	6,496551724	4,01753818
Fin de semana	0	hay_derrota_encajando_primero	461	5,735357918	3,701854314
Fin de semana	1	hay_derrota_encajando_primero	275	5,607272727	3,521355835
Entre semana	0	hay_derrota_local	358	6,773743017	4,466079182
Entre semana	1	hay_derrota_local	65	6,753846154	4,138352515
Fin de semana	0	hay_derrota_local	610	5,639344262	3,621831209
Fin de semana	1	hay_derrota_local	126	5,920634921	3,6956259
Entre semana	0	hay_derrota_visitante	323	6,866873065	4,514048307
Entre semana	1	hay_derrota_visitante	100	6,46	4,073615524
Fin de semana	0	hay_derrota_visitante	527	5,70398482	3,667714872
Fin de semana	1	hay_derrota_visitante	209	5,645933014	3,55438451

Test de diferencia de medias						
variable	estadistico	p_value	conf_low	conf_high	mensaje	tipo_dia
hay_futbol	2,131222438	0,033239527	0,020892569	0,504149148	OK	Total
hay_futbol	-0,290251364	0,771753549	-0,495569661	0,368014733	OK	Entre semana
hay_futbol	-2,846742035	0,00449612	-0,721611261	-0,132758078	OK	Fin de semana
hay_derrota	0,771646544	0,440488529	-0,277143091	0,636425913	OK	Total
hay_derrota	0,745557592	0,456396628	-0,523258677	1,1624412	OK	Entre semana
hay_derrota	0,044752499	0,964317162	-0,515457388	0,539504646	OK	Fin de semana
hay_victoria	-1,058560189	0,29007455	-0,728920101	0,218104575	OK	Total
hay_victoria	0,519894904	0,60344574	-0,632774972	1,087646131	OK	Entre semana
hay_victoria	-1,677554523	0,094002226	-1,022536001	0,080507016	OK	Fin de semana
hay_partido_tension_alta	-2,270444154	0,024430273	-1,698917028	-0,118676834	OK	Total
hay_partido_tension_alta	-1,211376848	0,229704025	-2,299743844	0,561176351	OK	Entre semana
hay_partido_tension_alta	-1,811710694	0,073058322	-1,714508152	0,07790041	OK	Fin de semana
hay_derrota_tension_alta	-1,228014036	0,222406044	-1,460078345	0,343877227	OK	Total
hay_derrota_tension_alta	-0,15326932	0,878841829	-1,627371901	1,396954147	OK	Entre semana
hay_derrota_tension_alta	-1,29241491	0,202141313	-1,817346764	0,394224075	OK	Fin de semana
hay_derrota_descuento	-1,272489439	0,208648369	-1,995508342	0,445922443	OK	Total
hay_derrota_descuento	-1,017017075	0,319812948	-3,353559497	1,143629776	OK	Entre semana
hay_derrota_descuento	-0,593872427	0,556976765	-1,761807402	0,967582166	OK	Fin de semana
hay_derrota_tardia	0,494732721	0,621320561	-0,488896795	0,816435426	OK	Total
hay_derrota_tardia	-0,243694479	0,808191785	-1,417560757	1,108942705	OK	Entre semana
hay_derrota_tardia	0,843069466	0,400675101	-0,417816189	1,038761746	OK	Fin de semana
hay_remontada_sufrida	-0,759085982	0,449670895	-1,225679426	0,547599075	OK	Total
hay_remontada_sufrida	-0,495605288	0,62480774	-2,490381932	1,527108298	OK	Entre semana
hay_remontada_sufrida	-0,906055963	0,367979402	-1,433923375	0,537942804	OK	Fin de semana
hay_derrota_abultada	0,133382207	0,894139499	-0,743832875	0,851160596	OK	Total
hay_derrota_abultada	1,407037813	0,167311605	-0,394216546	2,196718191	OK	Entre semana
hay_derrota_abultada	-0,838037739	0,404970127	-1,460345332	0,596614515	OK	Fin de semana
hay_derbi	-2,465761079	0,023573432	-4,100789552	-0,332543781	OK	Total
hay_derbi	-0,688256599	0,521500366	-8,278178491	4,769785205	OK	Entre semana
hay_derbi	-3,494108049	0,00400241	-4,064285733	-0,957099521	OK	Fin de semana
hay Eliminacion	-0,674898113	0,504080425	-2,018135328	1,010509502	OK	Total
hay Eliminacion	-0,411683641	0,684039582	-2,518717272	1,679118275	OK	Entre semana
hay Eliminacion	0,512677927	0,618722856	-1,395699385	2,237831046	OK	Fin de semana
hay_derrota_con_racha_previa	0,112397863	0,910750226	-0,811562463	0,908944298	OK	Total
hay_derrota_con_racha_previa	-0,947916431	0,349751746	-2,527895123	0,919106447	OK	Entre semana
hay_derrota_con_racha_previa	1,511426437	0,135826277	-0,202538026	1,456971677	OK	Fin de semana
hay_derrota_encajando_primerero	1,119872932	0,263053123	-0,1988892	0,727557284	OK	Total
hay_derrota_encajando_primerero	0,962984719	0,336262925	-0,434978843	1,269213525	OK	Entre semana
hay_derrota_encajando_primerero	0,46827325	0,639759477	-0,409102111	0,665272491	OK	Fin de semana
hay_derrota_local	-0,472628988	0,636851264	-0,750524757	0,459916389	OK	Total
hay_derrota_local	0,03521756	0,971981593	-1,101992639	1,141786365	OK	Entre semana
hay_derrota_local	-0,780464908	0,436153459	-0,992523054	0,429941737	OK	Fin de semana
hay_derrota_visitante	0,930369263	0,352562387	-0,262749477	0,735743956	OK	Total
hay_derrota_visitante	0,850185548	0,396347907	-0,537435747	1,351181877	OK	Entre semana
hay_derrota_visitante	0,197985001	0,843159156	-0,518410509	0,63451412	OK	Fin de semana

Matriz de correlaciones por franjas



9.4 Script del Análisis en RStudio

```
#####  
# ANÁLISIS ESTADÍSTICO Y ECONOMETRICO EN RSTUDIO  
#####  
  
rm(list = ls())  
  
options(repos = c(CRAN = "https://cloud.r-project.org"))  
  
setwd("C:/Users/yagom/OneDrive/Documentos/TFG ADE DEF")  
  
getwd()  
list.files()  
  
paquetes <- c(  
  "readr",  
  "dplyr",  
  "tidyr",  
  "lubridate",  
  "janitor",  
  "ggplot2",  
  "corrplot",  
  "psych",  
  "broom",
```

```

    "lmtest",
    "sandwich",
    "car",
    "MASS",
    "pROC",
    "randomForest",
    "stringr"
)

paquetes_instalados <- rownames(installed.packages())

paquetes_faltantes <- paquetes[!(paquetes %in% paquetes_instalados)]

if(length(paquetes_faltantes) > 0){
  install.packages(paquetes_faltantes, dependencies = TRUE)
}

library(readr)
library(dplyr)
library(tidyr)
library(lubridate)
library(janitor)
library(ggplot2)
library(corrplot)
library(psych)
library(broom)
library(lmtest)
library(sandwich)
library(car)
library(MASS)
library(pROC)
library(randomForest)
library(stringr)

dir.create("outputs", showWarnings = FALSE)
dir.create("outputs/tablas", showWarnings = FALSE, recursive = TRUE)
dir.create("outputs/graficos", showWarnings = FALSE, recursive = TRUE)
dir.create("outputs/modelos", showWarnings = FALSE, recursive = TRUE)

#####
# 1. FUNCIONES AUXILIARES
#####

rename_first <- function(df, new_name, candidates){

  existing <- candidates[candidates %in% names(df)]

  if(length(existing) > 0){

    old_name <- existing[1]

    if(!(new_name %in% names(df))){

```

```

    names(df)[names(df) == old_name] <- new_name
  }
}

return(df)
}

ensure_numeric <- function(df, var){

  if(!(var %in% names(df))){
    df[[var]] <- 0
  }

  df[[var]] <- suppressWarnings(as.numeric(df[[var]]))
  df[[var]][is.na(df[[var])] <- 0

  return(df)
}

safe_t_test <- function(data, y, group){

  data_aux <- data %>%
    filter(!is.na(.data[[y]]), !is.na(.data[[group]]))

  if(length(unique(data_aux[[group]])) < 2){

    return(data.frame(
      variable = group,
      estadistico = NA,
      p_value = NA,
      conf_low = NA,
      conf_high = NA,
      mensaje = "No hay dos grupos suficientes"
    ))
  }

  formula_test <- as.formula(paste(y, "~", group))

  test <- tryCatch(
    t.test(formula_test, data = data_aux),
    error = function(e) NULL
  )

  if(is.null(test)){

    return(data.frame(
      variable = group,
      estadistico = NA,
      p_value = NA,
      conf_low = NA,
      conf_high = NA,

```

```

    mensaje = "Error en t.test"
  ))
}

return(data.frame(
  variable = group,
  estadistico = as.numeric(test$statistic),
  p_value = test$p.value,
  conf_low = test$conf.int[1],
  conf_high = test$conf.int[2],
  mensaje = "OK"
))
}

robust_lm_table <- function(modelo, nombre_modelo){

  robust <- coefest(modelo, vcov = vcovHC(modelo, type = "HC1"))

  tabla <- data.frame(
    modelo = nombre_modelo,
    variable = rownames(robust),
    estimate = robust[, 1],
    std_error_robust = robust[, 2],
    statistic = robust[, 3],
    p_value = robust[, 4],
    row.names = NULL
  )

  return(tabla)
}

safe_lm <- function(formula, data){

  tryCatch(
    lm(formula, data = data),
    error = function(e){
      message("Error en modelo LM: ", e$message)
      return(NULL)
    }
  )
}

safe_glm <- function(formula, data, family){

  tryCatch(
    glm(formula, data = data, family = family),
    error = function(e){
      message("Error en modelo GLM: ", e$message)
      return(NULL)
    }
  )
}

```

```
}
```

```
safe_glm_nb <- function(formula, data){
```

```
  tryCatch(  
    MASS::glm.nb(formula, data = data),  
    error = function(e){  
      message("Error en modelo binomial negativa: ", e$message)  
      return(NULL)  
    }  
  )  
}
```

```
safe_vif <- function(modelo, nombre_modelo){
```

```
  resultado <- tryCatch(  
    {  
      v <- car::vif(modelo)  
  
      if(is.matrix(v)){  
  
        tabla <- data.frame(  
          modelo = nombre_modelo,  
          variable = rownames(v),  
          as.data.frame(v),  
          row.names = NULL  
        )  
  
      } else {  
  
        tabla <- data.frame(  
          modelo = nombre_modelo,  
          variable = names(v),  
          vif = as.numeric(v),  
          row.names = NULL  
        )  
      }  
  
      tabla  
    },  
    error = function(e){  
  
      data.frame(  
        modelo = nombre_modelo,  
        variable = NA,  
        vif = NA  
      )  
    }  
  )  
  
  return(resultado)
```

```

}

#####
# 2. CARGA DE DATOS
#####

linea_diario <- readLines("dataset_r_diario.csv", n = 1, encoding = "UTF-8")
linea_franjas <- readLines("dataset_r_franjas.csv", n = 1, encoding = "UTF-8")

cat("Cabecera diario:\n")
cat(linea_diario, "\n\n")

cat("Cabecera franjas:\n")
cat(linea_franjas, "\n\n")

cat("Separadores detectados en diario:\n")
cat("Comas:", str_count(linea_diario, ","), "\n")
cat("Punto y coma:", str_count(linea_diario, ";"), "\n")
cat("Tabuladores:", str_count(linea_diario, "\t"), "\n\n")

cat("Separadores detectados en franjas:\n")
cat("Comas:", str_count(linea_franjas, ","), "\n")
cat("Punto y coma:", str_count(linea_franjas, ";"), "\n")
cat("Tabuladores:", str_count(linea_franjas, "\t"), "\n\n")

cargar_csv_auto <- function(archivo){

  linea <- readLines(archivo, n = 1, encoding = "UTF-8")

  n_comas <- str_count(linea, ",")
  n_punto_coma <- str_count(linea, ";")
  n_tabs <- str_count(linea, "\t")

  if(n_punto_coma >= n_comas & n_punto_coma >= n_tabs){

    datos <- read_delim(
      archivo,
      delim = ";",
      locale = locale(encoding = "UTF-8"),
      show_col_types = FALSE
    )

  } else if(n_tabs >= n_comas & n_tabs >= n_punto_coma){

    datos <- read_delim(
      archivo,
      delim = "\t",
      locale = locale(encoding = "UTF-8"),
      show_col_types = FALSE
    )

  } else {

```

```

    datos <- read_delim(
      archivo,
      delim = ",",
      locale = locale(encoding = "UTF-8"),
      show_col_types = FALSE
    )
  }

  return(datos)
}

diario <- cargar_csv_auto("dataset_r_diario.csv")
franjas <- cargar_csv_auto("dataset_r_franjas.csv")

diario <- diario %>%
  clean_names()

franjas <- franjas %>%
  clean_names()

cat("Columnas diario:", ncol(diario), "\n")
cat("Columnas franjas:", ncol(franjas), "\n\n")

print(names(diario))
print(names(franjas))

glimpse(diario)
glimpse(franjas)

c("fecha", "provincia", "llamadas", "tipo_dia") %in% names(diario)

#####
# 3. NORMALIZACIÓN DE NOMBRES
#####

diario <- diario %>%
  rename_first("fecha", c("fecha", "date", "key_date")) %>%
  rename_first("provincia", c("provincia", "province")) %>%
  rename_first("caa", c("caa", "comunidad_autonoma", "comunidad")) %>%
  rename_first("ano", c("ano", "anio", "año", "year")) %>%
  rename_first("mes", c("mes", "month")) %>%
  rename_first("mes_num", c("mes_num", "mes_numero", "n_mes", "numero_mes",
"month_num")) %>%
  rename_first("dia_semana", c("dia_semana", "dia_de_la_semana", "weekday"))
%>%
  rename_first("tipo_dia", c("tipo_dia", "tipo_de_dia")) %>%
  rename_first("llamadas", c("llamadas", "llamadas_pertinentes",
"total_llamadas")) %>%
  rename_first("partidos", c("partidos", "total_partidos", "n_partidos")) %>%
  rename_first("victorias", c("victorias", "total_victorias")) %>%
  rename_first("derrotas", c("derrotas", "total_derrotas")) %>%

```

```

rename_first("empates", c("empates", "total_empates")) %>%
rename_first("goles_equipo", c("goles_equipo", "goles_favor",
"goles_a_favor")) %>%
rename_first("goles_rival", c("goles_rival", "goles_contra",
"goles_en_contra")) %>%
rename_first("goles_totales", c("goles_totales", "total_goles")) %>%
rename_first("dif_goles", c("dif_goles", "diferencia_goles",
"balance_goles")) %>%
rename_first("derrota_tardia", c("derrota_tardia", "derrotas_tardias")) %>%
rename_first("derrota_descuento", c("derrota_descuento",
"derrotas_descuento")) %>%
rename_first("derrota_abultada", c("derrota_abultada", "derrotas_abultadas"))
%>%
rename_first("remontada_sufrida", c("remontada_sufrida",
"remontadas_sufridas")) %>%
rename_first("derbi", c("derbi", "derbis")) %>%
rename_first("eliminacion", c("eliminacion", "eliminaciones")) %>%
rename_first("ultimas_3_jornadas", c("ultimas_3_jornadas",
"ultimas_tres_jornadas")) %>%
rename_first("partido_nocturno", c("partido_nocturno", "partidos_nocturnos"))
%>%
rename_first("partido_alta_intensidad", c("partido_alta_intensidad",
"hay_alta_intensidad", "partidos_alta_intensidad")) %>%
rename_first("partido_tension_alta", c("partido_tension_alta",
"partidos_tension_alta")) %>%
rename_first("derrota_tension_alta", c("derrota_tension_alta",
"derrotas_tension_alta")) %>%
rename_first("racha_derrotas_previa_max", c("racha_derrotas_previa_max",
"racha_derrotas_previa", "racha_previa")) %>%
rename_first("derrota_con_racha_previa", c("derrota_con_racha_previa",
"derrotas_con_racha_previa")) %>%
rename_first("gol_contra_temprano", c("gol_contra_temprano",
"goles_contra_temprano")) %>%
rename_first("derrota_gol_contra_temprano", c("derrota_gol_contra_temprano",
"derrotas_gol_contra_temprano")) %>%
rename_first("encaja_primerero", c("encaja_primerero")) %>%
rename_first("derrota_encajando_primerero", c("derrota_encajando_primerero",
"derrotas_encajando_primerero")) %>%
rename_first("derrota_local", c("derrota_local", "derrotas_local")) %>%
rename_first("derrota_visitante", c("derrota_visitante",
"derrotas_visitante"))

```

```

#####
# 4. NORMALIZACIÓN DE NOMBRES - FRANJAS
#####

```

```

franjas <- franjas %>%
rename_first("provincia", c("provincia", "province")) %>%
rename_first("ccaa", c("ccaa", "comunidad_autonoma", "comunidad")) %>%
rename_first("ano", c("ano", "anio", "año", "year")) %>%
rename_first("mes", c("mes", "month")) %>%
rename_first("mes_num", c("mes_num", "mes_numero", "n_mes", "numero_mes",
"month_num")) %>%

```

```

rename_first("mes_ano", c("mes_ano")) %>%
rename_first("franja", c("franja", "franja_horaria", "franja_cruce",
"franja_1h")) %>%
rename_first("llamadas", c("llamadas", "llamadas_pertinentes",
"total_llamadas")) %>%
rename_first("partidos_durante", c("partidos_durante",
"total_partidos_durante")) %>%
rename_first("victorias_durante", c("victorias_durante",
"total_victorias_durante")) %>%
rename_first("derrotas_durante", c("derrotas_durante",
"total_derrotas_durante")) %>%
rename_first("partidos_posterior", c("partidos_posterior",
"partidos_franja_posterior", "total_partidos_posterior")) %>%
rename_first("victorias_posterior", c("victorias_posterior",
"victorias_franja_posterior")) %>%
rename_first("derrotas_posterior", c("derrotas_posterior",
"derrotas_franja_posterior")) %>%
rename_first("goles_totales_franja", c("goles_totales_franja",
"goles_totales_franja_gol")) %>%
rename_first("goles_contra_franja", c("goles_contra_franja",
"goles_contra_franja_gol", "goles_rival_franja")) %>%
rename_first("goles_favor_franja", c("goles_favor_franja",
"goles_favor_franja_gol", "goles_equipo_franja"))

print(names(diario))
print(names(franjas))

```

```

#####
# 5. PREPARACIÓN BASE DIARIA
#####

```

```

if("fecha" %in% names(diario)){

diario$fecha <- as.character(diario$fecha)

diario$fecha <- case_when(
  grepl("^\\d{8}$", diario$fecha) ~ as.character(as.Date(diario$fecha, format
= "%Y%m%d")),
  TRUE ~ diario$fecha
)

diario$fecha <- as.Date(diario$fecha)
}

```

```

vars_num_diario <- c(
  "llamadas",
  "partidos",
  "victorias",
  "derrotas",
  "empates",
  "goles_equipo",
  "goles_rival",
  "goles_totales",

```

```

"dif_goles",
"derrota_tardia",
"derrota_descuento",
"derrota_abultada",
"remontada_sufrida",
"derbi",
"eliminacion",
"ultimas_3_jornadas",
"partido_nocturno",
"partido_alta_intensidad",
"partido_tension_alta",
"derrota_tension_alta",
"racha_derrotas_previa_max",
"derrota_con_racha_previa",
"gol_contra_temprano",
"derrota_gol_contra_temprano",
"encaja_primer",
"derrota_encajando_primer",
"derrota_local",
"derrota_visitante"
)

for(v in vars_num_diario){
  diario <- ensure_numeric(diario, v)
}

diario <- diario %>%
  mutate(
    goles_totales = if_else(
      goles_totales == 0,
      goles_equipo + goles_rival,
      goles_totales
    ),
    dif_goles = goles_equipo - goles_rival
  )

#####
# 6. VARIABLES TEMPORALES Y BINARIAS - BASE DIARIA
#####

if(!("ano" %in% names(diario)) & "fecha" %in% names(diario)){
  diario$ano <- year(diario$fecha)
}

if(!("mes_num" %in% names(diario)) & "fecha" %in% names(diario)){
  diario$mes_num <- month(diario$fecha)
}

if(!("dia_semana" %in% names(diario)) & "fecha" %in% names(diario)){
  diario$dia_semana <- weekdays(diario$fecha)
}

if(!("tipo_dia" %in% names(diario)) & "fecha" %in% names(diario)){

```

```

diario$tipo_dia <- if_else(
  wday(diario$fecha, week_start = 1) %in% c(6, 7),
  "Fin de semana",
  "Entre semana"
)
}

diario <- diario %>%
  mutate(
    hay_futbol = if_else(partidos > 0, 1, 0),
    hay_derrota = if_else(derrotas > 0, 1, 0),
    hay_victoria = if_else(victorias > 0, 1, 0),
    hay_empate = if_else(empates > 0, 1, 0),
    hay_derrota_tardia = if_else(derrota_tardia > 0, 1, 0),
    hay_derrota_descuento = if_else(derrota_descuento > 0, 1, 0),
    hay_derrota_abultada = if_else(derrota_abultada > 0, 1, 0),
    hay_remontada_sufrida = if_else(remontada_sufrida > 0, 1, 0),
    hay_derbi = if_else(derbi > 0, 1, 0),
    hay Eliminacion = if_else(eliminacion > 0, 1, 0),
    hay_partido_tension_alta = if_else(partido_tension_alta > 0, 1, 0),
    hay_derrota_tension_alta = if_else(derrota_tension_alta > 0, 1, 0),
    hay_partido_alta_intensidad = if_else(partido_alta_intensidad > 0, 1, 0),
    hay_derrota_con_racha_previa = if_else(derrota_con_racha_previa > 0, 1, 0),
    hay_derrota_encajando_primerero = if_else(derrota_encajando_primerero > 0, 1,
0),
    hay_derrota_local = if_else(derrota_local > 0, 1, 0),
    hay_derrota_visitante = if_else(derrota_visitante > 0, 1, 0),
    fin_semana = if_else(tipo_dia == "Fin de semana", 1, 0),
    entre_semana = if_else(tipo_dia == "Entre semana", 1, 0)
  )

```

```

#####
# 7. CONTROLES TEMPORALES, LAGS Y DUPLICADOS
#####

```

```

diario <- diario %>%
  mutate(
    provincia = as.factor(provincia),
    tipo_dia = as.factor(tipo_dia),
    ano = as.factor(ano),
    mes_num = as.factor(mes_num),
    dia_semana = as.factor(dia_semana),
    mes_ano = interaction(ano, mes_num, drop = TRUE)
  ) %>%
  arrange(provincia, fecha) %>%
  group_by(provincia) %>%
  mutate(
    tendencia = as.numeric(fecha - min(fecha, na.rm = TRUE)) + 1,
    hay_futbol_lag1 = lag(hay_futbol, 1),
    hay_derrota_lag1 = lag(hay_derrota, 1),
    hay_derrota_descuento_lag1 = lag(hay_derrota_descuento, 1),
    hay_remontada_sufrida_lag1 = lag(hay_remontada_sufrida, 1),

```

```

    hay_partido_tension_alta_lag1 = lag(hay_partido_tension_alta, 1)
  ) %>%
  ungroup()

diario <- diario %>%
  mutate(
    hay_futbol_lag1 = if_else(is.na(hay_futbol_lag1), 0, hay_futbol_lag1),
    hay_derrota_lag1 = if_else(is.na(hay_derrota_lag1), 0, hay_derrota_lag1),
    hay_derrota_descuento_lag1 = if_else(is.na(hay_derrota_descuento_lag1), 0,
hay_derrota_descuento_lag1),
    hay_remontada_sufrida_lag1 = if_else(is.na(hay_remontada_sufrida_lag1), 0,
hay_remontada_sufrida_lag1),
    hay_partido_tension_alta_lag1 =
if_else(is.na(hay_partido_tension_alta_lag1), 0, hay_partido_tension_alta_lag1)
  )

duplicados_diario <- diario %>%
  count(provincia, fecha) %>%
  filter(n > 1)

write_csv(duplicados_diario, "outputs/tablas/duplicados_diario.csv")
write_csv(diario, "outputs/tablas/diario_preparado.csv")

print(duplicados_diario)

```

```

#####
# 8. PREPARACIÓN BASE FRANJAS
#####

```

```

vars_num_franjas <- c(
  "llamadas",
  "partidos_durante",
  "victorias_durante",
  "derrotas_durante",
  "partidos_posterior",
  "victorias_posterior",
  "derrotas_posterior",
  "goles_totales_franja",
  "goles_contra_franja",
  "goles_favor_franja"
)

for(v in vars_num_franjas){
  franjas <- ensure_numeric(franjas, v)
}

if(!("ano" %in% names(franjas))){
  franjas$ano <- NA
}

if(!("mes_num" %in% names(franjas))){
  franjas$mes_num <- NA
}

```

```

if(!("franja" %in% names(franjas))){
  franjas$franja <- NA
}

if(!("mes_ano" %in% names(franjas))){
  franjas$mes_ano <- interaction(franjas$ano, franjas$mes_num, drop = TRUE)
}

franjas <- franjas %>%
  mutate(
    provincia = as.factor(provincia),
    ano = as.factor(ano),
    mes_num = as.factor(mes_num),
    franja = as.factor(franja),
    mes_ano = as.factor(mes_ano),
    hay_partidos_durante = if_else(partidos_durante > 0, 1, 0),
    hay_partidos_posterior = if_else(partidos_posterior > 0, 1, 0),
    hay_goles_contra_franja = if_else(goles_contra_franja > 0, 1, 0)
  )

duplicados_franjas <- franjas %>%
  count(provincia, ano, mes_num, franja) %>%
  filter(n > 1)

write_csv(duplicados_franjas, "outputs/tablas/duplicados_franjas.csv")
write_csv(franjas, "outputs/tablas/franjas_preparado.csv")

print(duplicados_franjas)

#####
# 9. RESUMEN PREPARACIÓN BASES
#####

resumen_preparacion <- data.frame(
  base = c("diario", "franjas"),
  observaciones = c(nrow(diario), nrow(franjas)),
  columnas = c(ncol(diario), ncol(franjas)),
  duplicados_clave = c(nrow(duplicados_diario), nrow(duplicados_franjas))
)

write_csv(resumen_preparacion, "outputs/tablas/resumen_preparacion_bases.csv")

print(resumen_preparacion)

#####
# 10. ESTADÍSTICOS DESCRIPTIVOS
#####

descriptivos_llamadas <- diario %>%
  summarise(
    observaciones = n(),
    llamadas_media = mean(llamadas, na.rm = TRUE),

```

```

    llamadas_mediana = median(llamadas, na.rm = TRUE),
    llamadas_sd = sd(llamadas, na.rm = TRUE),
    llamadas_min = min(llamadas, na.rm = TRUE),
    llamadas_max = max(llamadas, na.rm = TRUE),
    total_llamadas = sum(llamadas, na.rm = TRUE)
  )

descriptivos_provincia <- diario %>%
  group_by(provincia) %>%
  summarise(
    observaciones = n(),
    llamadas_media = mean(llamadas, na.rm = TRUE),
    llamadas_mediana = median(llamadas, na.rm = TRUE),
    llamadas_sd = sd(llamadas, na.rm = TRUE),
    llamadas_min = min(llamadas, na.rm = TRUE),
    llamadas_max = max(llamadas, na.rm = TRUE),
    total_llamadas = sum(llamadas, na.rm = TRUE),
    partidos_total = sum(partidos, na.rm = TRUE),
    derrotas_total = sum(derrotas, na.rm = TRUE),
    victorias_total = sum(victorias, na.rm = TRUE),
    .groups = "drop"
  ) %>%
  arrange(desc(total_llamadas))

descriptivos_tipo_dia <- diario %>%
  group_by(tipo_dia) %>%
  summarise(
    observaciones = n(),
    llamadas_media = mean(llamadas, na.rm = TRUE),
    llamadas_mediana = median(llamadas, na.rm = TRUE),
    llamadas_sd = sd(llamadas, na.rm = TRUE),
    total_llamadas = sum(llamadas, na.rm = TRUE),
    partidos_media = mean(partidos, na.rm = TRUE),
    derrotas_media = mean(derrotas, na.rm = TRUE),
    .groups = "drop"
  )

descriptivos_futbol <- diario %>%
  summarise(
    partidos_media = mean(partidos, na.rm = TRUE),
    derrotas_media = mean(derrotas, na.rm = TRUE),
    victorias_media = mean(victorias, na.rm = TRUE),
    empates_media = mean(empates, na.rm = TRUE),
    goles_rival_media = mean(goles_rival, na.rm = TRUE),
    goles_equipo_media = mean(goles_equipo, na.rm = TRUE),
    dif_goles_media = mean(dif_goles, na.rm = TRUE),
    dias_con_futbol = sum(hay_futbol, na.rm = TRUE),
    dias_con_derrota = sum(hay_derrota, na.rm = TRUE),
    dias_con_derrota_descuento = sum(hay_derrota_descuento, na.rm = TRUE),
    dias_con_remontada = sum(hay_remontada_sufrida, na.rm = TRUE),
    dias_con_derbi = sum(hay_derbi, na.rm = TRUE),
    dias_con_tension_alta = sum(hay_partido_tension_alta, na.rm = TRUE)
  )

```

```

write_csv(descriptivos_llamadas,
"outputs/tablas/descriptivos_llamadas_general.csv")
write_csv(descriptivos_provincia,
"outputs/tablas/descriptivos_por_provincia.csv")
write_csv(descriptivos_tipo_dia,
"outputs/tablas/descriptivos_por_tipo_dia.csv")
write_csv(descriptivos_futbol,
"outputs/tablas/descriptivos_variables_futbol.csv")

print(descriptivos_llamadas)
print(descriptivos_provincia)
print(descriptivos_tipo_dia)
print(descriptivos_futbol)

#####
# 11. GRÁFICOS DESCRIPTIVOS BÁSICOS
#####
options(install.packages.compile.from.source="never")

install.packages(
  "ggplot2",
  type="binary",
  repos="https://cloud.r-project.org"
)

library(ggplot2)

graf_llamadas_provincia <- ggplot(diario, aes(x = provincia, y = llamadas)) +
  geom_boxplot() +
  labs(
    title = "Distribución de llamadas al 016 por provincia",
    x = "Provincia",
    y = "Llamadas"
  ) +
  theme_minimal()

ggsave(
  "outputs/graficos/boxplot_llamadas_provincia.png",
  graf_llamadas_provincia,
  width = 10,
  height = 6
)

graf_llamadas_tipo_dia <- ggplot(diario, aes(x = tipo_dia, y = llamadas)) +
  geom_boxplot() +
  labs(
    title = "Distribución de llamadas al 016 por tipo de día",
    x = "Tipo de día",
    y = "Llamadas"
  ) +
  theme_minimal()

```

```

ggsave(
  "outputs/graficos/boxplot_llamadas_tipo_dia.png",
  graf_llamadas_tipo_dia,
  width = 8,
  height = 6
)
print(graf_llamadas_provincia)
print(graf_llamadas_tipo_dia)

#####
# 12. MATRIZ DE CORRELACIONES COMPLETA
#####
library(dplyr)
library(readr)
library(corrplot)

vars_cor <- diario %>%
  dplyr::select(
    llamadas,
    partidos,
    victorias,
    derrotas,
    empates,
    goles_equipo,
    goles_rival,
    goles_totales,
    dif_goles,
    derrota_tension_alta,
    derrota_descuento,
    derrota_tardia,
    derrota_abultada,
    remontada_sufrida,
    derbi,
    eliminacion,
    partido_tension_alta,
    partido_alta_intensidad,
    racha_derrotas_previa_max
  )
matriz_cor <- cor(
  vars_cor,
  use = "pairwise.complete.obs",
  method = "pearson"
)

write.csv(
  round(matriz_cor, 4),
  "outputs/tablas/matriz_correlaciones_completa.csv",
  row.names = TRUE,
  fileEncoding = "UTF-8"
)

cor_llamadas <- data.frame(
  variable = names(matriz_cor[, "llamadas"]),

```

```

    correlacion_con_llamadas = as.numeric(matriz_cor[, "llamadas"])
  ) %>%
  arrange(desc(correlacion_con_llamadas))

write_csv(cor_llamadas, "outputs/tablas/correlaciones_con_llamadas.csv")

png("outputs/graficos/matriz_correlaciones_completa.png", width = 1200, height
= 900)

corrplot(
  matriz_cor,
  method = "color",
  type = "upper",
  tl.cex = 0.7,
  number.cex = 0.6,
  addCoef.col = "black"
)

dev.off()

print(cor_llamadas)

#####
# 13. MATRIZ DE CORRELACIONES REDUCIDA
#####

vars_cor_reducida <- diario %>%
  dplyr::select(
    llamadas,
    partidos,
    derrotas,
    victorias,
    goles_rival,
    dif_goles,
    derrota_tension_alta,
    derrota_descuento,
    remontada_sufrida,
    derbi,
    partido_tension_alta
  )

matriz_cor_reducida <- cor(
  vars_cor_reducida,
  use = "pairwise.complete.obs",
  method = "pearson"
)

write.csv(
  round(matriz_cor_reducida, 4),
  "outputs/tablas/matriz_correlaciones_reducida.csv",
  row.names = TRUE,
  fileEncoding = "UTF-8"
)

```



```

    TRUE,
    TRUE,
    TRUE
  )
)

tabla_medias_total <- data.frame()
tabla_tests_total <- data.frame()

for(i in 1:nrow(eventos_comparacion)){

  var_evento <- eventos_comparacion$variable[i]
  solo_futbol <- eventos_comparacion$solo_dias_futbol[i]

  data_base <- if(solo_futbol) diario_futbol else diario

  medias <- data_base %>%
    group_by(tipo_dia, valor_evento = .data[[var_evento]]) %>%
    summarise(
      variable = var_evento,
      n = n(),
      llamadas_media = mean(llamadas, na.rm = TRUE),
      llamadas_sd = sd(llamadas, na.rm = TRUE),
      .groups = "drop"
    )

  tabla_medias_total <- bind_rows(tabla_medias_total, medias)

  test_general <- safe_t_test(data_base, "llamadas", var_evento) %>%
    mutate(tipo_dia = "Total")

  test_entre <- safe_t_test(data_base %>% filter(tipo_dia == "Entre semana"),
"llamadas", var_evento) %>%
    mutate(tipo_dia = "Entre semana")

  test_finde <- safe_t_test(data_base %>% filter(tipo_dia == "Fin de semana"),
"llamadas", var_evento) %>%
    mutate(tipo_dia = "Fin de semana")

  tabla_tests_total <- bind_rows(tabla_tests_total, test_general, test_entre,
test_finde)
}

write_csv(tabla_medias_total, "outputs/tablas/comparacion_medias_eventos.csv")
write_csv(tabla_tests_total, "outputs/tablas/tests_medias_eventos.csv")

print(tabla_medias_total)
print(tabla_tests_total)

```

```

#####
# 15. MODELOS LINEALES SOBRE LA BASE DIARIA
#####

```

```

modelo_1 <- safe_lm(
  llamadas ~ hay_futbol + fin_semana + factor(provincia) + factor(mes_ano),
  diario
)

modelo_2 <- safe_lm(
  llamadas ~ hay_derrota + hay_victoria + goles_rival + dif_goles + fin_semana
+ factor(provincia) + factor(mes_ano),
  diario_futbol
)

modelo_3 <- safe_lm(
  llamadas ~ hay_derrota_descuento + hay_remontada_sufrida + hay_derbi +
hay_partido_tension_alta + fin_semana + factor(provincia) + factor(mes_ano),
  diario_futbol
)

modelo_4 <- safe_lm(
  llamadas ~ hay_derrota_tension_alta + hay_derrota_descuento +
hay_remontada_sufrida + hay_derrota_abultada + fin_semana + factor(provincia) +
factor(mes_ano),
  diario_futbol
)

modelo_5_entre <- safe_lm(
  llamadas ~ hay_futbol + hay_derrota + hay_derrota_descuento +
hay_remontada_sufrida + hay_partido_tension_alta + factor(provincia) +
factor(mes_ano),
  diario %>% filter(tipo_dia == "Entre semana")
)

modelo_6_finde <- safe_lm(
  llamadas ~ hay_futbol + hay_derrota + hay_derrota_descuento +
hay_remontada_sufrida + hay_partido_tension_alta + factor(provincia) +
factor(mes_ano),
  diario %>% filter(tipo_dia == "Fin de semana")
)

modelo_7_lags <- safe_lm(
  llamadas ~ hay_futbol + hay_futbol_lag1 + hay_derrota_lag1 +
hay_derrota_descuento_lag1 + hay_remontada_sufrida_lag1 + fin_semana +
factor(provincia) + factor(mes_ano),
  diario
)

#####
# 16. TABLAS MODELOS LINEALES CON ERRORES ROBUSTOS
#####
options(install.packages.compile.from.source="never")

install.packages("lmtest", type="binary", repos="https://cloud.r-project.org")
install.packages("sandwich", type="binary", repos="https://cloud.r-project.org")

```

```

library(lmtest)
library(sandwich)
tablas_lm <- list()

if(!is.null(modelo_1)){
  tablas_lm[["modelo_1_actividad_futbol"]] <- robust_lm_table(modelo_1, "Modelo
1 - Actividad futbolística")
}

if(!is.null(modelo_2)){
  tablas_lm[["modelo_2_resultado_general"]] <- robust_lm_table(modelo_2,
"Modelo 2 - Resultado general")
}

if(!is.null(modelo_3)){
  tablas_lm[["modelo_3_intensidad_emocional"]] <- robust_lm_table(modelo_3,
"Modelo 3 - Intensidad emocional")
}

if(!is.null(modelo_4)){
  tablas_lm[["modelo_4_eventos_negativos"]] <- robust_lm_table(modelo_4,
"Modelo 4 - Eventos negativos")
}

if(!is.null(modelo_5_entre)){
  tablas_lm[["modelo_5_entre_semana"]] <- robust_lm_table(modelo_5_entre,
"Modelo 5 - Entre semana")
}

if(!is.null(modelo_6_finde)){
  tablas_lm[["modelo_6_fin_semana"]] <- robust_lm_table(modelo_6_finde, "Modelo
6 - Fin de semana")
}

if(!is.null(modelo_7_lags)){
  tablas_lm[["modelo_7_lags"]] <- robust_lm_table(modelo_7_lags, "Modelo 7 -
Retardos")
}

tabla_lm_total <- bind_rows(tablas_lm)

write_csv(tabla_lm_total, "outputs/modelos/modelos_lineales_robustos.csv")

tabla_lm_variables_clave <- tabla_lm_total %>%
  filter(
    variable != "(Intercept)",
    !grepl("factor\\(provincia\\)", variable),
    !grepl("factor\\(mes_ano\\)", variable)
  )

write_csv(tabla_lm_variables_clave,
"outputs/modelos/modelos_lineales_variables_clave.csv")

```

```
print(tabla_lm_variables_clave)
```

```
#####  
# 17. MULTICOLINEALIDAD - VIF  
#####
```

```
vif_modelos <- data.frame()
```

```
if(!is.null(modelo_2)){  
  vif_modelos <- bind_rows(  
    vif_modelos,  
    safe_vif(modelo_2, "Modelo 2 - Resultado general")  
  )  
}
```

```
if(!is.null(modelo_3)){  
  vif_modelos <- bind_rows(  
    vif_modelos,  
    safe_vif(modelo_3, "Modelo 3 - Intensidad emocional")  
  )  
}
```

```
if(!is.null(modelo_4)){  
  vif_modelos <- bind_rows(  
    vif_modelos,  
    safe_vif(modelo_4, "Modelo 4 - Eventos negativos")  
  )  
}
```

```
write_csv(vif_modelos, "outputs/modelos/vif_modelos.csv")
```

```
print(vif_modelos)
```

```
#####  
# 18. MODELOS POISSON PARA DATOS DE CONTEO  
#####
```

```
poisson_1 <- safe_glm(  
  llamadas ~ hay_futbol + fin_semana + factor(provincia) + factor(mes_ano),  
  diario,  
  poisson(link = "log")  
)
```

```
poisson_2 <- safe_glm(  
  llamadas ~ hay_derrota + hay_victoria + goles_rival + dif_goles + fin_semana  
+ factor(provincia) + factor(mes_ano),  
  diario_futbol,  
  poisson(link = "log")  
)
```

```
poisson_3 <- safe_glm(  
  llamadas ~ hay_derrota + hay_victoria + goles_rival + dif_goles + fin_semana  
+ factor(provincia) + factor(mes_ano),  
  diario_futbol,  
  poisson(link = "log")  
)
```

```

    llamadas ~ hay_derrota_descuento + hay_remontada_sufrida + hay_derbi +
    hay_partido_tension_alta + fin_semana + factor(provincia) + factor(mes_ano),
    diario_futbol,
    poisson(link = "log")
)

```

```
sobredispersion <- data.frame()
```

```
if(!is.null(poisson_1)){
```

```

    sobredispersion <- bind_rows(
      sobredispersion,
      data.frame(
        modelo = "Poisson 1 - Actividad futbolística",
        dispersion = sum(residuals(poisson_1, type = "pearson")^2) /
poisson_1$df.residual
      )
    )
}

```

```
if(!is.null(poisson_2)){
```

```

    sobredispersion <- bind_rows(
      sobredispersion,
      data.frame(
        modelo = "Poisson 2 - Resultado general",
        dispersion = sum(residuals(poisson_2, type = "pearson")^2) /
poisson_2$df.residual
      )
    )
}

```

```
if(!is.null(poisson_3)){
```

```

    sobredispersion <- bind_rows(
      sobredispersion,
      data.frame(
        modelo = "Poisson 3 - Intensidad emocional",
        dispersion = sum(residuals(poisson_3, type = "pearson")^2) /
poisson_3$df.residual
      )
    )
}

```

```
write_csv(sobredispersion, "outputs/modelos/sobredispersion_poisson.csv")
```

```
print(sobredispersion)
```

```

#####
# 19. MODELOS BINOMIAL NEGATIVA
#####

```

```
nb_1 <- safe_glm_nb(
```

```

    llamadas ~ hay_futbol + fin_semana + factor(provincia) + factor(mes_ano),
    diario
  )

nb_2 <- safe_glm_nb(
  llamadas ~ hay_derrota + hay_victoria + goles_rival + dif_goles + fin_semana
+ factor(provincia) + factor(mes_ano),
  diario_futbol
)

nb_3 <- safe_glm_nb(
  llamadas ~ hay_derrota_descuento + hay_remontada_sufrida + hay_derbi +
hay_partido_tension_alta + fin_semana + factor(provincia) + factor(mes_ano),
  diario_futbol
)

#####
# 20. TABLAS MODELOS DE CONTEO
#####

tablas_count <- list()

if(!is.null(poisson_1)){

  tablas_count[["poisson_1"]] <- broom::tidy(poisson_1, exponentiate = TRUE,
conf.int = FALSE) %>%
  mutate(modelo = "Poisson 1 - Actividad futbolística")
}

if(!is.null(poisson_2)){

  tablas_count[["poisson_2"]] <- broom::tidy(poisson_2, exponentiate = TRUE,
conf.int = FALSE) %>%
  mutate(modelo = "Poisson 2 - Resultado general")
}

if(!is.null(poisson_3)){

  tablas_count[["poisson_3"]] <- broom::tidy(poisson_3, exponentiate = TRUE,
conf.int = FALSE) %>%
  mutate(modelo = "Poisson 3 - Intensidad emocional")
}

if(!is.null(nb_1)){

  tablas_count[["nb_1"]] <- broom::tidy(nb_1, exponentiate = TRUE, conf.int =
FALSE) %>%
  mutate(modelo = "Binomial negativa 1 - Actividad futbolística")
}

if(!is.null(nb_2)){

```

```

  tablas_count[["nb_2"]] <- broom::tidy(nb_2, exponentiate = TRUE, conf.int =
FALSE) %>%
  mutate(modelo = "Binomial negativa 2 - Resultado general")
}

if(!is.null(nb_3)){

  tablas_count[["nb_3"]] <- broom::tidy(nb_3, exponentiate = TRUE, conf.int =
FALSE) %>%
  mutate(modelo = "Binomial negativa 3 - Intensidad emocional")
}

tabla_count_total <- bind_rows(tablas_count)

write_csv(tabla_count_total, "outputs/modelos/modelos_conteo_irr.csv")

tabla_count_variables_clave <- tabla_count_total %>%
  filter(
    term != "(Intercept)",
    !grepl("factor\\(provincia\\)", term),
    !grepl("factor\\(mes_ano\\)", term)
  )

write_csv(tabla_count_variables_clave,
"outputs/modelos/modelos_conteo_variables_clave.csv")

print(tabla_count_variables_clave)

#####
# 21. ANÁLISIS EXPLORATORIO POR FRANJAS HORARIAS
#####

descriptivos_franjas <- franjas %>%
  summarise(
    observaciones = n(),
    llamadas_media = mean(llamadas, na.rm = TRUE),
    llamadas_mediana = median(llamadas, na.rm = TRUE),
    llamadas_sd = sd(llamadas, na.rm = TRUE),
    llamadas_min = min(llamadas, na.rm = TRUE),
    llamadas_max = max(llamadas, na.rm = TRUE),
    partidos_durante_total = sum(partidos_durante, na.rm = TRUE),
    partidos_posterior_total = sum(partidos_posterior, na.rm = TRUE),
    goles_contra_franja_total = sum(goles_contra_franja, na.rm = TRUE)
  )

descriptivos_franja_horaria <- franjas %>%
  group_by(franja) %>%
  summarise(
    observaciones = n(),
    llamadas_media = mean(llamadas, na.rm = TRUE),
    llamadas_total = sum(llamadas, na.rm = TRUE),
    partidos_durante_total = sum(partidos_durante, na.rm = TRUE),
    partidos_posterior_total = sum(partidos_posterior, na.rm = TRUE),

```

```

    goles_contra_franja_total = sum(goles_contra_franja, na.rm = TRUE),
    .groups = "drop"
  )

write_csv(descriptivos_franjas,
"outputs/tablas/descriptivos_franjas_general.csv")
write_csv(descriptivos_franja_horaria,
"outputs/tablas/descriptivos_por_franja.csv")

print(descriptivos_franjas)
print(descriptivos_franja_horaria)

#####
# 22. MATRIZ DE CORRELACIONES POR FRANJAS
#####

vars_cor_franjas <- franjas %>%
  dplyr::select(
    llamadas,
    partidos_durante,
    partidos_posterior,
    goles_totales_franja,
    goles_contra_franja,
    goles_favor_franja,
    derrotas_durante,
    derrotas_posterior
  )

matriz_cor_franjas <- cor(
  vars_cor_franjas,
  use = "pairwise.complete.obs",
  method = "pearson"
)

write.csv(
  round(matriz_cor_franjas, 4),
  "outputs/tablas/matriz_correlaciones_franjas.csv",
  row.names = TRUE,
  fileEncoding = "UTF-8"
)

png("outputs/graficos/matriz_correlaciones_franjas.png", width = 900, height =
700)

corrplot(
  matriz_cor_franjas,
  method = "color",
  type = "upper",
  tl.cex = 0.8,
  number.cex = 0.7,
  addCoef.col = "black"
)

```

```

dev.off()

print(matriz_cor_franjas)

#####
# 23. MODELOS EXPLORATORIOS POR FRANJAS
#####

modelo_franjas_1 <- safe_lm(
  llamadas ~ partidos_durante + partidos_posterior + goles_contra_franja +
  factor(provincia) + factor(mes_ano),
  franjas
)

modelo_franjas_2 <- safe_lm(
  llamadas ~ hay_partidos_durante + hay_partidos_posterior +
  hay_goles_contra_franja + factor(provincia) + factor(mes_ano),
  franjas
)

tablas_franjas <- list()

if(!is.null(modelo_franjas_1)){
  tablas_franjas[["modelo_franjas_1"]] <- robust_lm_table(modelo_franjas_1,
  "Franjas 1 - Variables continuas")
}

if(!is.null(modelo_franjas_2)){
  tablas_franjas[["modelo_franjas_2"]] <- robust_lm_table(modelo_franjas_2,
  "Franjas 2 - Variables binarias")
}

tabla_franjas_modelos <- bind_rows(tablas_franjas)

write_csv(tabla_franjas_modelos,
"outputs/modelos/modelos_franjas_robustos.csv")

tabla_franjas_variables_clave <- tabla_franjas_modelos %>%
  filter(
    variable != "(Intercept)",
    !grepl("factor\\(provincia\\)", variable),
    !grepl("factor\\(mes_ano\\)", variable)
  )

write_csv(tabla_franjas_variables_clave,
"outputs/modelos/modelos_franjas_variables_clave.csv")

print(tabla_franjas_variables_clave)

#####
# 24. GRÁFICO LLAMADAS POR FRANJA
#####

```

```

graf_franjas_llamadas <- ggplot(descriptivos_franja_horaria, aes(x = franja, y
= llamadas_media, group = 1)) +
  geom_line() +
  geom_point() +
  labs(
    title = "Media de llamadas al 016 por franja horaria",
    x = "Franja horaria",
    y = "Media de llamadas"
  ) +
  theme_minimal() +
  theme(axis.text.x = element_text(angle = 45, hjust = 1))

ggsave(
  "outputs/graficos/llamadas_media_por_franja.png",
  graf_franjas_llamadas,
  width = 12,
  height = 6
)

#####
# 25. ANÁLISIS PREDICTIVO EXPLORATORIO - LOGIT
#####

diario_pred <- diario %>%
  group_by(provincia) %>%
  mutate(
    p75_provincia = quantile(llamadas, 0.75, na.rm = TRUE),
    llamadas_altas = if_else(llamadas > p75_provincia, 1, 0)
  ) %>%
  ungroup()

set.seed(123)

indices_train <- sample(
  1:nrow(diario_pred),
  size = floor(0.7 * nrow(diario_pred))
)

train <- diario_pred[indices_train, ]
test <- diario_pred[-indices_train, ]

logit_1 <- safe_glm(
  llamadas_altas ~ hay_futbol + hay_derrota + hay_victoria + goles_rival +
  dif_goles + hay_derrota_descuento + hay_remontada_sufrida + hay_derbi +
  hay_partido_tension_alta + fin_semana + factor(provincia) + factor(mes_num),
  train,
  binomial(link = "logit")
)

tabla_logit <- data.frame()

if(!is.null(logit_1)){

```

```

tabla_logit <- broom::tidy(logit_1, exponentiate = TRUE, conf.int = FALSE)
%>%
  mutate(modelo = "Logit - llamadas altas")

pred_logit <- predict(logit_1, newdata = test, type = "response")

auc_logit <- tryCatch(
  as.numeric(pROC::auc(test$llamadas_altas, pred_logit)),
  error = function(e) NA
)

metricas_logit <- data.frame(
  modelo = "Logit",
  auc = auc_logit
)

write_csv(tabla_logit, "outputs/modelos/modelo_logit_odds_ratios.csv")
write_csv(metricas_logit, "outputs/modelos/metricas_logit.csv")
}

tabla_logit_variables_clave <- tabla_logit %>%
  filter(
    term != "(Intercept)",
    !grepl("factor\\(provincia\\)", term),
    !grepl("factor\\(mes_num\\)", term)
  )

write_csv(tabla_logit_variables_clave,
"outputs/modelos/modelo_logit_variables_clave.csv")

print(tabla_logit_variables_clave)

#####
# 26. RANDOM FOREST EXPLORATORIO
#####
options(install.packages.compile.from.source="never")

install.packages(
  "randomForest",
  type="binary",
  repos="https://cloud.r-project.org"
)

library(randomForest)
vars_rf <- c(
  "llamadas_altas",
  "hay_futbol",
  "hay_derrota",
  "hay_victoria",
  "goles_rival",
  "dif_goles",
  "hay_derrota_descuento",

```

```

    "hay_remontada_sufrida",
    "hay_derbi",
    "hay_partido_tension_alta",
    "fin_semana",
    "provincia",
    "mes_num"
  )

train_rf <- train %>%
  dplyr::select(all_of(vars_rf)) %>%
  mutate(
    llamadas_altas = as.factor(llamadas_altas),
    provincia = as.factor(provincia),
    mes_num = as.factor(mes_num)
  )

test_rf <- test %>%
  dplyr::select(all_of(vars_rf)) %>%
  mutate(
    llamadas_altas = as.factor(llamadas_altas),
    provincia = as.factor(provincia),
    mes_num = as.factor(mes_num)
  )

rf_1 <- tryCatch(
  randomForest(
    llamadas_altas ~ .,
    data = train_rf,
    ntree = 500,
    importance = TRUE
  ),
  error = function(e){
    message("Error en Random Forest: ", e$message)
    return(NULL)
  }
)

if(!is.null(rf_1)){

  pred_rf_prob <- predict(rf_1, newdata = test_rf, type = "prob")[, 2]

  auc_rf <- tryCatch(
    as.numeric(pROC::auc(as.numeric(as.character(test_rf$llamadas_altas)),
pred_rf_prob)),
    error = function(e) NA
  )

  importancia_rf <- data.frame(
    variable = rownames(importance(rf_1)),
    importance(rf_1),
    row.names = NULL
  )
}

```

```

metricas_rf <- data.frame(
  modelo = "Random Forest",
  auc = auc_rf
)

write_csv(importancia_rf,
"outputs/modelos/random_forest_importancia_variables.csv")
write_csv(metricas_rf, "outputs/modelos/metricas_random_forest.csv")

png("outputs/graficos/random_forest_importancia.png", width = 900, height =
700)

varImpPlot(rf_1)

dev.off()

print(importancia_rf)
print(metricas_rf)
}

#####
# 27. RESUMEN FINAL DE ARCHIVOS GENERADOS
#####

resumen_outputs <- data.frame(
  archivo = list.files("outputs", recursive = TRUE),
  carpeta = "outputs"
)

write_csv(resumen_outputs, "outputs/resumen_archivos_generados.csv")

print("SCRIPT FINALIZADO")
print("Revisa las carpetas outputs/tablas, outputs/graficos y outputs/modelos")

print(resumen_outputs)

```

9.5 Declaración de uso de herramientas de inteligencia artificial generativa

Declaración de Uso de Herramientas de Inteligencia Artificial Generativa en Trabajos Fin de Grado

ADVERTENCIA: Desde la Universidad consideramos que ChatGPT u otras herramientas similares son herramientas muy útiles en la vida académica, aunque su uso queda siempre bajo la responsabilidad del alumno, puesto que las respuestas que proporciona pueden no ser veraces. En este sentido, NO está permitido su uso en la elaboración del Trabajo fin de Grado para generar código porque estas herramientas no son fiables en esa tarea. Aunque el código funcione, no hay garantías de que metodológicamente sea correcto, y es altamente probable que no lo sea.

Por la presente, yo, [Yago Maus Rumbao], estudiante de [5E2 + Analytics] de la Universidad Pontificia Comillas al presentar mi Trabajo Fin de Grado titulado "[FÚTBOL Y VIOLENCIA DE GÉNERO EN ANDALUCÍA Y LA REGIÓN DE MURCIA]", declaro que he utilizado la herramienta de Inteligencia Artificial Generativa ChatGPT u otras similares de IAG de código sólo en el contexto de las actividades descritas a continuación [el alumno debe mantener solo aquellas en las que se ha usado ChatGPT o similares y borrar el resto. Si no se ha usado ninguna, borrar todas y escribir "no he usado ninguna"]:

1. **Referencias:** Usado conjuntamente con otras herramientas, como Science, para identificar referencias preliminares que luego he contrastado y validado.
2. **Metodólogo:** Para descubrir métodos aplicables a problemas específicos de investigación.
3. **Interpretador de código:** Para realizar análisis de datos preliminares.
4. **Corrector de estilo literario y de lenguaje:** Para mejorar la calidad lingüística y estilística del texto.
5. **Sintetizador y divulgador de libros complicados:** Para resumir y comprender literatura compleja.
6. **Revisor:** Para recibir sugerencias sobre cómo mejorar y perfeccionar el trabajo con diferentes niveles de exigencia.

Afirmo que toda la información y contenido presentados en este trabajo son producto de mi investigación y esfuerzo individual, excepto donde se ha indicado lo contrario y se han dado los créditos correspondientes (he incluido las referencias adecuadas en el TFG y he explicitado para que se ha usado ChatGPT u otras herramientas similares). Soy consciente de las implicaciones académicas y éticas de presentar un trabajo no original y acepto las consecuencias de cualquier violación a esta declaración.

Fecha: [3/6/2026]

Firma: _____ Yago Maus _____