METODOLOGÍA Y EL INFORME COMPLETO

Cómo se construyó esto

Cada número de la página principal sale de un solo conjunto de datos versionado y del código de este repositorio. Aquí está qué es ese conjunto de datos, cómo una pausa se convierte en un número de momentum, qué controlan las comparaciones y dónde están los límites honestos.

---

00 — LOS HALLAZGOS, EN BREVE

El equipo que va arriba en el momentum pierde alrededor de **−25** puntos en los cinco minutos después de una pausa de hidratación obligatoria. Pero los *mismos* equipos pierden alrededor de **−17** en minutos tranquilos, sin pausa — así que la mayor parte de la caída es regresión a la media, el enfriamiento natural tras un buen rato, no el silbatazo.

Lo que queda es una diferencia de unos **8** puntos entre pausa y no pausa para los mismos equipos. Eso es compatible con un efecto real pequeño, pero los intervalos se traslapan y la muestra todavía es chica, así que es sugerente, no comprobado. Una mirada aparte al clima sugiere que la mayoría de los partidos nunca llegó al calor para el que sirve una pausa de enfriamiento. Nada de esto es todavía un veredicto causal — la eliminatoria lo afinará.

01 — QUÉ SE MIDE AQUÍ

La pregunta es acotada y comprobable: ¿las pausas de hidratación obligatorias de la FIFA mueven el momentum en contra del equipo que iba arriba? La variable es el índice de momentum por minuto de **FotMob** — su modelo de amenaza esperada sobre qué lado domina, armado con el flujo de ataques y ocasiones, no con el marcador. Nosotros *leemos* ese número; no calculamos el nuestro. Positivo es que el local aprieta, negativo el visitante.

Es un análisis vivo: la página se reconstruye desde el conjunto de datos versionado cada jornada hasta la final del 19 de julio, así que las cifras se mueven conforme entran datos.

02 — DE DÓNDE SALEN LOS DATOS

Tres fuentes, cada una con un trabajo:

- **FotMob** — la serie de momentum por minuto y las alineaciones (que se usan en la prueba de calor).
- **ESPN** — la narración de texto, que da el momento exacto de las pausas (y de eventos de VAR / lesión), así que las pausas se detectan por lo que de verdad pasó, no por un reloj fijo.
- **Open-Meteo** — temperatura y humedad en cada sede y horario de inicio para el WBGT, y el clima histórico de las ciudades de los clubes para la prueba de aclimatación.

Publicamos solo datos *derivados*: el conjunto de datos procesado y los resúmenes con fecha quedan versionados en el repo, pero los datos crudos nunca se redistribuyen, por respeto a los términos de las fuentes. El scraping corre localmente en una conexión de casa; el sitio público lo reconstruye CI desde el conjunto de datos versionado y nunca hace scraping. El historial de git es el sistema de instantáneas — la estimación de cada día es un commit.

03 — DE UNA PAUSA A UN NÚMERO

Por cada pausa detectada se toman dos ventanas de cinco minutos de momentum: la ventana **previa** (los cinco minutos antes del silbatazo) y la **posterior** (los cinco minutos después), excluyendo el minuto mismo de la pausa. La variable es el cambio entre ambas — `momentum_delta = media posterior − media previa`.

El momentum es local-en-positivo, así que cada pausa produce *dos* filas, una desde la perspectiva de cada equipo (la del visitante es solo el negativo). Como se reflejan, el promedio agrupado es cero por construcción. Por eso siempre reportamos al equipo que iba **arriba** antes de la pausa — la acusación de "matar el momentum" es justamente que la pausa empuja el momentum lejos de quien iba ganándolo.

04 — LAS REFERENCIAS DE COMPARACIÓN

Un equipo que acaba de tener cinco minutos ardientes tiende a enfriarse de todos modos, con pausa o sin ella. Esa regresión a la media es la mayor amenaza para leer la caída cruda como un efecto. Así que corremos la *misma medición exacta* donde no había pausa obligatoria, en la misma escala de FotMob, y comparamos:

- **Los mismos equipos de 2026 en minutos tranquilos** (−17) — el control más limpio: mismos equipos, mismo torneo, solo que sin silbatazo. Todo lo fijo de un equipo —su nivel, su calor, su altitud— se cancela aquí.
- **Mundial 2022** (−15) y **Euro 2024** (−16) — futbol de selecciones en climas más frescos, sin pausas obligatorias.
- **Copa América 2024** (−10) — selecciones en el calor del verano de Estados Unidos, pero una sola edición más ruidosa.
- **Mundial de Clubes 2025** (−23) — clubes, que regresan más fuerte que las selecciones; un contraste, no una comparación pareja.

El futbol de selecciones regresa por su cuenta alrededor de −10 a −17. La pausa (−25) queda unos 8 puntos por debajo del control de mismos equipos — la parte que la regresión a la media no explica.

05 — INTERVALOS DE CONFIANZA

Varias pausas dentro de un mismo partido no son independientes, así que los intervalos se calculan por bootstrap remuestreando *partidos*, no filas (un bootstrap por conglomerados). Temprano en el torneo hay pocos conglomerados, así que el intervalo del 95% es ancho — léelo como indicativo, no como un valor p preciso. El efecto se sostiene con ventanas de 4, 5 o 6 minutos. Por ahora el titular descansa en 86 pausas de hidratación con equipo arriba, y la gráfica de la estimación a lo largo del torneo, en la página principal, muestra si se está estabilizando o desvaneciendo.

06 — LA PRUEBA DE CALOR Y ACLIMATACIÓN

La objeción intuitiva es que los vaivenes son por el calor, no por el silbatazo — jugadores de ligas frescas derritiéndose en el verano de Estados Unidos. Lo probamos directo. Para cada equipo armamos una *brecha de aclimatación*: el WBGT del día del partido menos el WBGT al que está acostumbrado el plantel en casa, ligando a cada jugador con la ciudad de su club (391 clubes ubicados). Si el calor por desplazamiento moviera las caídas, los equipos más lejos de casa deberían caer más fuerte.

No lo hacen. Entre torneos, la mayor brecha va con la *menor* caída, no con la mayor:

Torneo	brecha	caída
Copa América 2024	+11°C	-10
Mundial de Clubes 2025	+8°C	-23
Mundial 2026	+5°C	-19
Euro 2024	+4°C	-16

Dentro de cada grupo es la misma historia. Entre los clubes del Mundial de Clubes la pendiente brecha→caída es +0.99 por °C [-0.30, +2.29] — si acaso, los clubes más lejos de casa cayeron *menos*. Agrupando las selecciones es -0.12 por °C [-0.84, +0.55], prácticamente plana. Así que las caídas grandes no son aclimatación, y la cifra alta del Mundial de Clubes es estructural —los clubes regresan más— no calor.

Un matiz honesto: la brecha de calor está enredada con el continente, la liga y la carga de partidos, así que esto descarta la versión simple del calor más que probar un cero exacto. Y de todos modos nunca amenazó al titular — el control de mismos equipos ya deja fijo el calor de cada equipo.

07 — ALTITUD

La altitud es un factor distinto del calor —aire delgado, no termorregulación— y solo dos sedes son altas (Ciudad de México y Guadalajara, ambas por encima de los 1,500 m). Son muy pocas para probarlo, y una pausa de enfriamiento no es para eso, así que lo anotamos y lo dejamos.

08 — LO QUE SE PUEDE Y NO SE PUEDE DECIR

- **Muestra chica, por ahora.** Temprano en el torneo los intervalos son anchos; la gráfica de la estimación a lo largo del tiempo muestra si el efecto se estabiliza o se desvanece.
- **Duraciones de pausa, ahora medidas.** Las marcas de inicio/fin de ESPN dan una duración exacta para 86 de 86 pausas de hidratación con equipo arriba (mediana 2:51, 1:36–3:56); la cobertura de VAR y lesión es más delgada. Si las pausas *más largas* pegan más fuerte es inconcluso con la muestra actual —el intervalo de la pendiente incluye el cero— así que no nos apoyamos en eso.
- **El WBGT es una estimación a la sombra** a partir de temperatura y humedad; el calor real en la cancha bajo el sol es mayor.
- **La brecha de aclimatación es colineal** con el continente, la liga y el calendario, así que es sugerente, no un instrumento limpio.
- **Todavía no hay titular causal.** El modelo causal acordado —una regresión de efectos fijos a dos vías, con errores por conglomerado de partido y la interacción hidratación×momentum-previo— se mantiene en pausa hasta que la muestra viva sea lo bastante grande para estimaciones estables.

09 — REPRODUCIBILIDAD

Cada cifra de aquí y de la página principal se calcula de forma determinista desde un solo archivo —el conjunto de datos procesado y versionado— con el código del repositorio. Nada se captura a mano. El código completo, los datos y este informe son públicos.

github.com/valternunez/wc2026-momentum ↗