

Creación de curvas de referencia Lichess estratificadas para Elo+Chess

Notas de investigación Elo+Chess

31 de mayo de 2026

Abstract

Los informes de referencia Elo+Chess se crean a partir de grandes muestras estratificadas de juegos Lichess.org, con la muestra de referencia utilizada en esta nota que contiene 483.974 partidas bala 1+0, 479.974 partidas blitz 3+0, 925.636 partidas rápidas 10+0, y 660.540 partidas rápidas más largas. Para cada tipo de juego importante, la cartera se expande cada juego en los estados del tablero antes y después del movimiento, calcula cientos de variables intermedias y del juego del jugador, agrega esas variables en métricas de entrenamiento y luego promedia esas métricas mediante el cubo Lichess Elo. La muestra está deliberadamente estratificada en lugar de ser proporcional al jugador original. pool: el objetivo es estimar curvas de comportamiento estable en cada nivel de calificación, incluyendo cubos que de otro modo serían escasos, para no reproducir la natural distribución de calificaciones. El resultado es una familia de curvas de referencia empíricas: por cada métrica, podemos preguntar con qué frecuencia los jugadores en diferentes niveles de Elo muestran un determinado Hábito o principio en los juegos reales. Esta nota describe el método, la escala de el cálculo, y varios ejemplos de principios de apertura que involucran enroque, desarrollo de piezas menores, movimientos repetidos de la pieza inicial y movimiento temprano de la reina. El énfasis es metodológico: Elo+Chess no asume que los principios del ajedrez sean universalmente lineal en todos los niveles de habilidad. En cambio, las curvas de referencia son inspeccionadas por rango de calificación, y el producto se enfoca en la región desde principiante hasta avanzado temprano donde las relaciones son más fuertes y más interpretables.

1 Objetivo

El objetivo del sistema de referencia es convertir los consejos generales de ajedrez en cantidades mensurables. Una afirmación como "desarrolla tus piezas menores" es útil, pero se vuelve más práctico cuando podemos decir cuántas piezas menores El jugador generalmente se desarrolla en el movimiento 10, ¿cómo se compara eso con sus compañeros en el mismo? grupo de calificación y cómo ese hábito cambia en todo el espectro de calificación.

Por lo tanto, Elo+Chess crea una biblioteca de métricas a partir de juegos reales. Cada métrica es calculado para cada fila de juego de jugador, resumido por tipo de juego y grupo de Elo, y utilizada como curva de referencia en el informe. El informe puede entonces colocar a un usuario en la misma curva y muestra si el usuario está por encima, por debajo o cerca del comportamiento observado entre jugadores comparables.

2 Datos de origen y estratificación

La fuente de referencia es la base de datos pública del juego Lichess.org. Usos de Elo+Chess grandes muestras estratificadas por tipo de juego y categoría de calificación en lugar de una pequeña muestra de conveniencia. Los artefactos de tiempo de ejecución de producción actuales cubren cuatro principales categorías de juegos:

- 1+0 bala,
- bombardeo 3+0,
- 10+0 rápido,
- Juegos rápidos de más de 10 minutos.

El objetivo de estratificación es la amplitud entre niveles de calificación. Un azar proporcional La muestra sería eficiente para describir el grupo general de jugadores, pero ser ineficiente para construir curvas de referencia. Los grupos de calificación más comunes dominaría la muestra, mientras que los cubos de cola de bajo soporte tendrían muy pocos juegos para promedios métricos estables. Elo+Chess por lo tanto muestra por juego promedio Cubo Elo. En la construcción actual de muestra limpia, los juegos sin procesar son lo primero filtrado a juegos clasificados válidos con jugadores conocidos, clasificaciones conocidas, resultados válidos, y controles de tiempo conocidos. Luego, los juegos se filtran a los jugadores que tienen suficientes juegos. en esa velocidad y calificaciones lo suficientemente estables para representar un nivel de habilidad coherente: Para cada jugador, Elo+Chess calcula el rango de clasificación dentro de la velocidad del jugador. $(\max(\text{Elo}) - \min(\text{Elo}))$ en todo el grupo de origen y mantiene juegos solo cuando ambos jugadores tienen al menos el recuento mínimo de juegos configurado y no más de 300 puntos Elo de deriva dentro de la velocidad. El filtro de deriva no está diseñado eliminar mejoras o variaciones normales; Reduce los casos en los que un jugador Los juegos combinarían niveles de habilidad sustancialmente diferentes, como cuentas provisionales, jugadores que regresan, o cuentas que se movieron rápidamente a través de varios grupos durante el período muestreado. Finalmente, los juegos elegibles se clasifican dentro de cada promedio-Elo. cubo por un hash determinista de la identificación y la semilla del juego, y un número limitado de Los juegos se retienen de cada grupo.

Este diseño hace que las curvas de referencia sean mucho más útiles. Cada cubo Elo aporta suficientes observaciones para estimar el valor promedio de cada métrica, y los raros cubos de alta o baja calificación no quedan ahogados por la forma natural de la distribución de jugadores activos. Las curvas de referencia utilizadas por Elo+Chess son luego restringido al rango más relevante para el producto, aproximadamente principiante a través del juego avanzado temprano. En los informes de escala Lichess, ese rango primario es 600–1600. Cuando un usuario ve el informe en la escala Chess.com, el los grupos de calificación correspondientes se asignan al mismo Lichess subyacente grupos de referencia utilizando el método de mapeo de calificación multiplataforma separado descrito en el artículo complementario: [Estimating Cross-Platform Chess Rating Mappings with Regresión modal](#).

3 Distribuciones de calificación de jugadores activos

Si miramos la distribución de las calificaciones Elo del jugador Lichess.org al final de marzo de 2026, por tipo de juego principal, vemos los siguientes recuentos y percentiles. Para esta vista de

distribución, cada jugador se cuenta una vez por tipo de juego exacto, usando la última calificación observada de ese jugador en el raw correspondiente filas de control de tiempo, después de requerir al menos cinco juegos en ese juego exacto tipo.¹

Game type	Players	Median	75th pct.	80th pct.	% ≤ 1600	Exact 1500	1500 share
1+0 bullet	1,221,146	1,479	1,858	1,951	58.9%	4,756	0.4%
3+0 blitz	1,144,045	1,542	1,845	1,912	54.9%	2,315	0.2%
10+0 rapid	1,679,401	1,381	1,691	1,765	68.3%	3,472	0.2%
>10 rapid	668,255	1,400	1,686	1,753	68.2%	7,969	1.2%

Table 1: Última distribución de calificación observada de jugador activo Lichess por exactitud tipo de juego después de requerir al menos cinco juegos en ese tipo de juego exacto.

Esta distribución no debe confundirse con la muestra de referencia estratificada. La distribución describe el grupo de jugadores activos Lichess observado en bruto escanear. Luego, la muestra de referencia se estratifica deliberadamente para que los cubos Elo tener suficiente soporte para curvas métricas estables.

Para asignar una calificación Elo Lichess.org para un tipo de juego en particular a un clasificación Chess.com equivalente, Elo+Chess utiliza el modal específico del tipo de juego ecuaciones de regresión estimadas en el compañero [Elo+Chess rating-mapping paper](#):

$$\begin{aligned}
\hat{R}_{\text{Chess.com,bullet}} &= -531.71 + 0.9871 R_{\text{Lichess}}, \\
\hat{R}_{\text{Chess.com,blitz}} &= -551.54 + 1.0853 R_{\text{Lichess}}, \\
\hat{R}_{\text{Chess.com,10minrapid}} &= -495.04 + 1.0741 R_{\text{Lichess}}, \\
\hat{R}_{\text{Chess.com,>10rapid}} &= -369.02 + 0.9374 R_{\text{Lichess}}.
\end{aligned}$$

Esas ecuaciones mapean valores de calificación, no rangos percentiles. Los rangos percentiles son definido dentro de un grupo de jugadores en particular, y los jugadores Chess.com y Lichess Las piscinas tienen diferentes formas. En la Tabla 2, para el Tipo de juego rápido de 10 minutos, la primera columna ofrece una pequeña muestra de posibles Chess.com Valores de calificación Elo. La segunda columna da el correspondiente Rango percentil de Chess.com mostrado en Chess.com al 31 de mayo de 2026. Un 71,7

La clasificación en Lichess.org para un rápido de 10 minutos corresponde a una clasificación Elo de Lichess de 1,644² y ese valor se muestra en el tercera columna. Si aplicamos la ecuación de mapeo Chess.com–Lichess, 1,644 en Lichess equivale aproximadamente a 1271 en Chess.com. Suponiendo el estimado La ecuación de mapeo es razonable, la brecha +522 entre la calificación Chess.com en La columna 1 y la calificación asignada en la columna 4 se pueden explicar por un gran número. de jugadores casuales en el grupo de jugadores Chess.com inflando el Chess.com rango percentil. En otras palabras, estar en el percentil 72 en Lichess.org es un logro mayor que estar en el percentil 72 en Chess.com, porque Los rangos percentiles no se asignan claramente de un sistema a otro. Elo real Los valores se pueden mapear utilizando ecuaciones estimadas a partir de puntuaciones del mismo jugador en los dos sistemas.

¹We apply a five-game minimum because lightly active accounts disproportionately remain at the 1500 starting/provisional anchor, creating an artificial spike in all-player latest-rating distributions. Figures 1–4 muestra las distribuciones con y sin este filtro.

²Calculated from the cumulative distribution function of the ZXQPROT3ZXQ 10+0 rapid latest-rating sample after requiring at least five games in that exact game type; see Table 1 and Figure 3.

Chess.com rating	Chess.com pct.	Same-pct. Lichess	Mapped equiv.	Gap
749	71.7%	1,644	1,271	+522
1,012	86.2%	1,871	1,515	+503
1,358	95.4%	2,095	1,755	+397
2,383	99.9%	2,557	2,251	-132

Table 2: Comparación ilustrativa entre el percentil rápido de 10 minutos Chess.com puntos y los mismos puntos percentiles en el rápido activo Lichess 10+0 distribución después de requerir al menos cinco juegos 10+0. “Chess.com asignado equiv.” aplica el actual Elo+Chess Conversión rápida de escala de calificación 10+0 a la calificación Lichess del mismo percentil. Se incluyen filas de la cola más alta para mostrar la distinción de escala, pero las filas ajustadas La línea de calificación es más confiable en el rango de entrenamiento principal.

El motivo de las lagunas en la Tabla 2 es que el mapeo estima calificaciones de fuerza de juego equivalentes, mientras que un percentil describe posición dentro de la propia población de una plataforma. Un jugador del percentil 71,7 en Por lo tanto, Chess.com no es automáticamente un jugador del percentil 71,7 en Lichess, y viceversa. Los números observados concuerdan con la opinión pública de Chess.com. grupo rápido que tiene una masa mucho mayor de jugadores casuales o de menor calificación, mientras que el El grupo Lichess de control exacto activo está más concentrado entre los repetidores experimentados. jugadores. El punto clave no es que una escala percentil sea “correcta” y la otro es “incorrecto”; es que el rango percentil no es portátil entre plataformas a menos que las distribuciones de los jugadores subyacentes tengan la misma forma.

Note on the 1500 spike. The visible jump near 1500 está presente en los datos en lugar de crearse mediante suavizado.

En lo último Distribución de calificaciones observada, 1500 calificaciones exactas representan el 13,1

jugadores bala, 5.6

17,6

anclaje inicial o provisional en el sistema de calificación.

El pico exacto de 1500 es en gran medida un artefacto de cuentas ligeramente activas. a verifique esto, volvimos a calcular las mismas distribuciones de calificación más recientes después de requerir cada jugador debe tener al menos cinco juegos del mismo tipo de juego. El pico casi desaparece bajo este filtro: las calificaciones exactas de 1500 caen del 5,6 al 17,6

jugadores en la distribución de todos los jugadores a 0.2–1.2

distribución reportada en la Tabla 1.

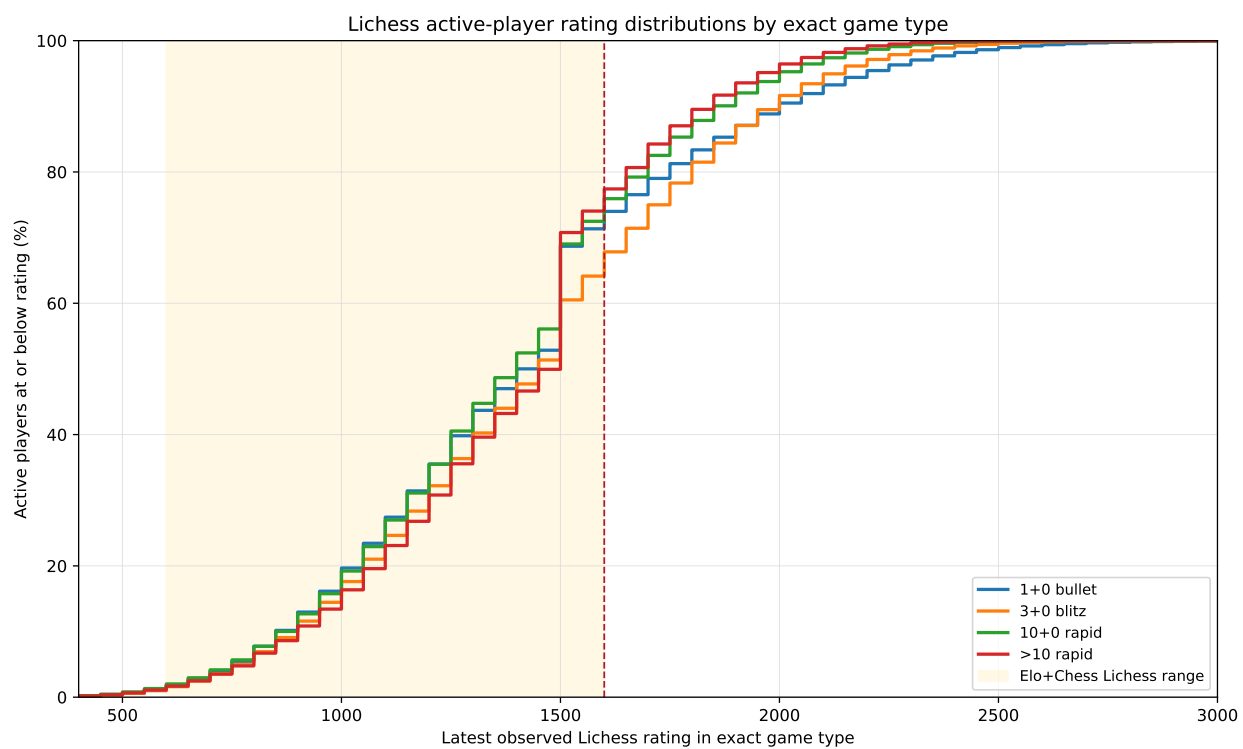


Figure 1: Distribuciones acumulativas de calificación de jugadores activos por tipo de juego exacto. el La banda amarilla marca el rango 600–1600 Lichess utilizado por el Elo+Chess principal. puntos de referencia del entrenamiento.

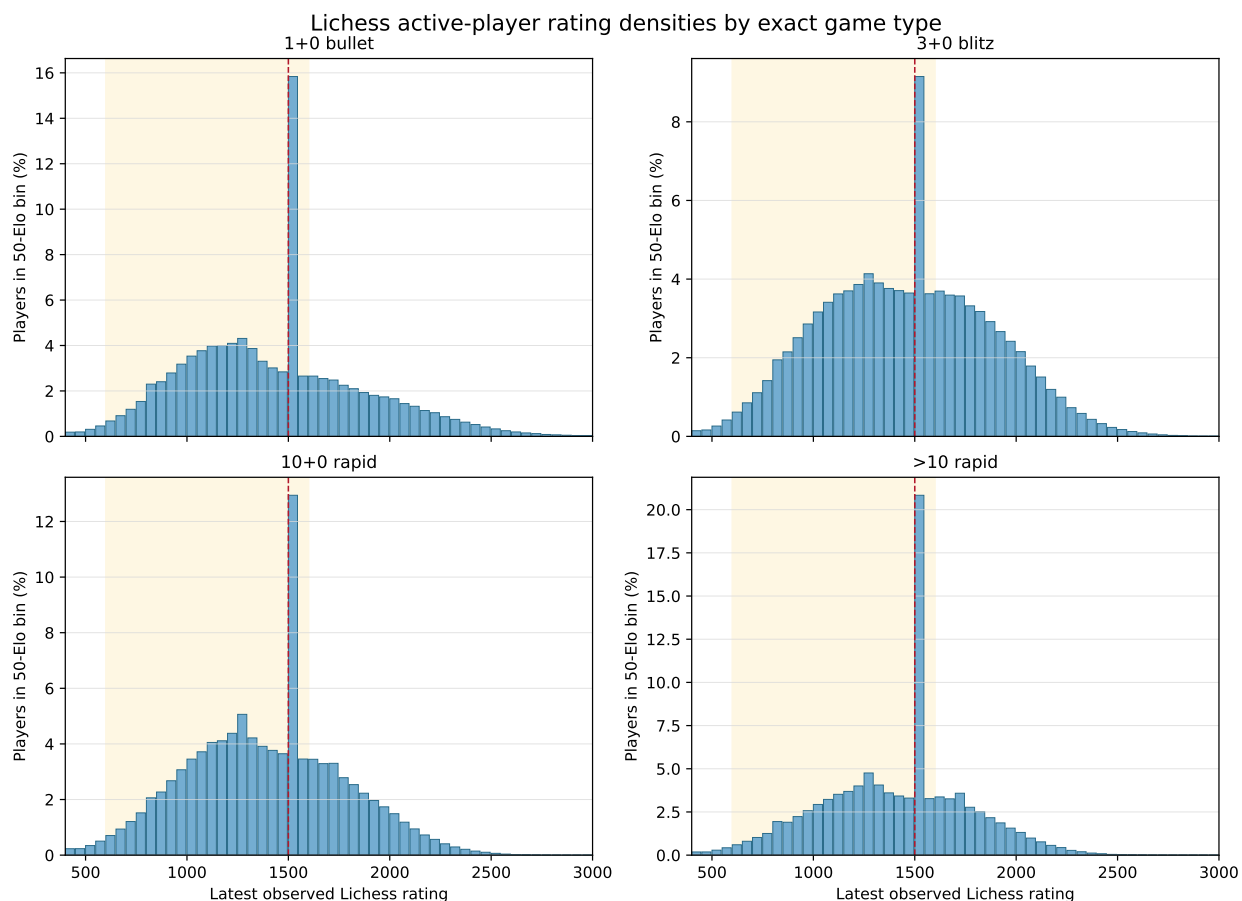


Figure 2: Vista de densidad de las mismas distribuciones de calificación de jugadores activos. el La línea discontinua marca 1500 Lichess, donde las distribuciones de calificación más recientes sin procesar tienen un gran punto de masa de clasificación exacta. La banda amarilla marca el 600–1600 Lichess rango utilizado por los principales informes de métricas Elo+Chess. El pico es más visible en la densidad de todos los jugadores porque muchas cuentas ligeramente activas permanecen exactamente en 1500; Figura 4 repite el gráfico de densidad después de requerir al menos cinco juegos en el tipo de juego exacto, lo que elimina en gran medida este artefacto.

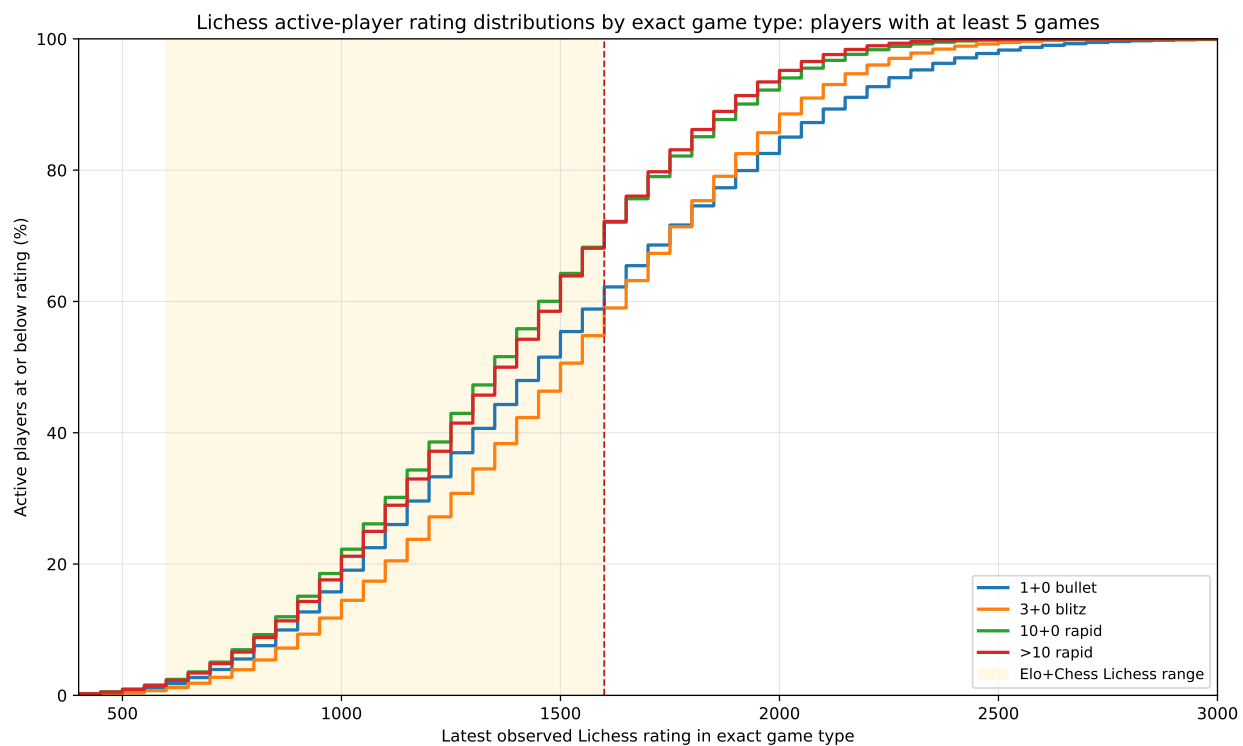


Figure 3: Distribuciones acumuladas de calificación de jugador activo después de requerir al menos al menos cinco juegos del mismo tipo de juego. El salto de 1500 se reduce mucho.

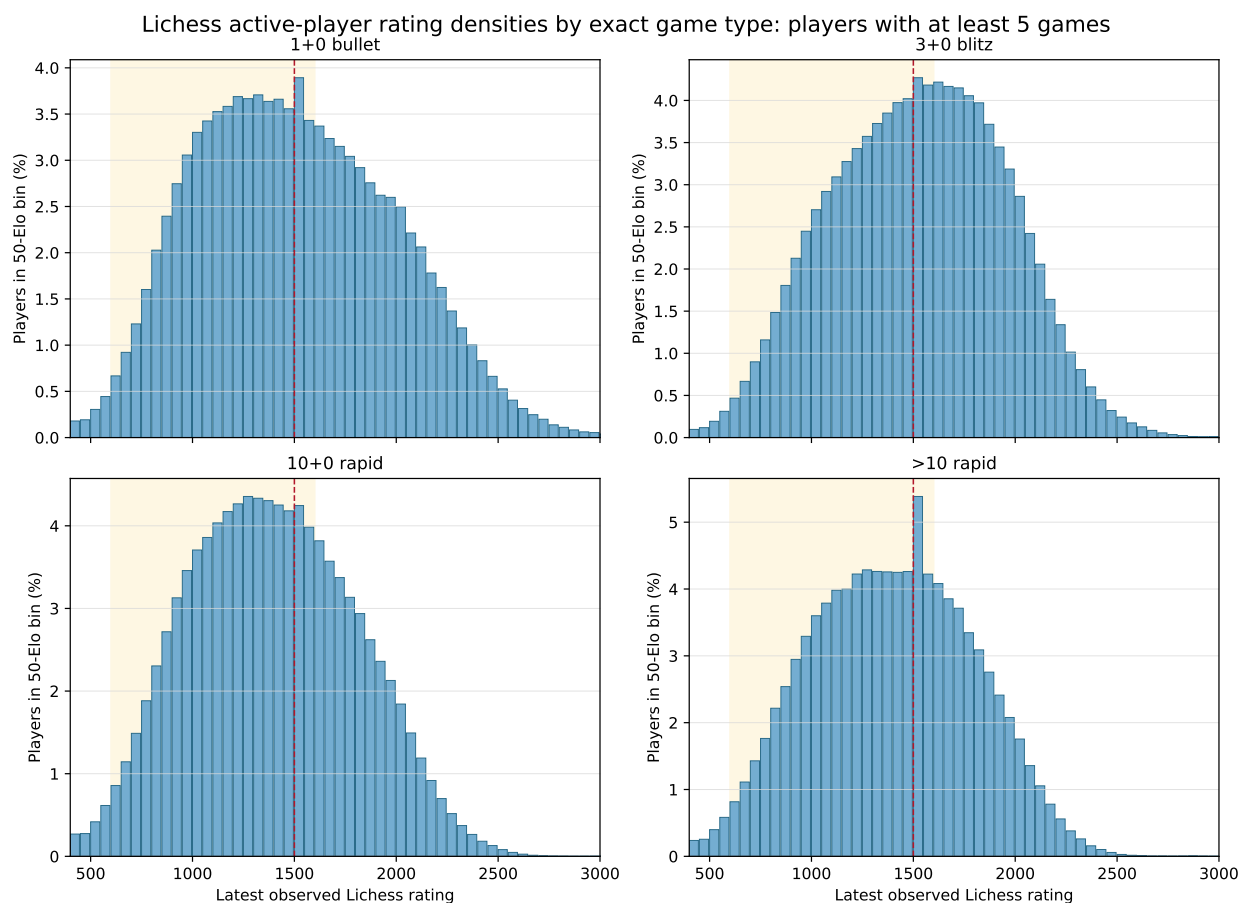


Figure 4: Vista de densidad después de requerir al menos cinco juegos en el juego exacto tipo. Esta vista confirma que el gran punto de masa de 1500 en el juego para todos los jugadores la distribución está impulsada principalmente por cuentas ligeramente activas. La banda amarilla marca el rango 600–1600 Lichess utilizado por los principales informes métricos Elo+Chess.

4 Escala de los artefactos de referencia actuales

El cálculo es intencionalmente grande. No es un raspado de perfil, un pequeño conjunto de PGN descargados o una colección cuidadosamente seleccionada de ejemplos instructivos. El pase inicial Lichess cubría archivos mensuales completos desde enero de 2025 hasta marzo de 2026. La tabla de escaneo no estratificada contiene 1,382,178,087 juegos y 2.764.356.174 filas de clasificación de juegos de jugadores. En esas filas, 6.623.438 únicos Los nombres de usuario de los jugadores aparecen al menos una vez.

Speed	Games in raw scan	Player-game rows	Unique players
Bullet	519,375,423	1,038,750,846	3,077,635
Blitz	839,823,863	1,679,647,726	5,987,634
Rapid	22,978,801	45,957,602	1,535,865
All scanned speeds	1,382,178,087	2,764,356,174	6,623,438

Table 3: Escala de escaneo Lichess no estratificada antes del muestreo de referencia. el escaneo Cubre archivos mensuales completos del historial del juego desde enero de 2025 hasta marzo. 2026.

Los artefactos de referencia de producción son mucho más pequeños que el escaneo sin procesar porque se seleccionan para la construcción métrica. Siguen siendo grandes: el punto de referencia La etapa es una tubería de expansión y agregación de estado de movimiento a lo largo de cientos de miles de juegos por tipo de juego.

Game type	Games	Player-game rows	Unique players	Move states	Feature columns
1+0 bullet	483,974	967,948	119,732	22,985,272	277
3+0 blitz	479,974	959,948	129,398	28,138,904	277
10+0 rapid	925,636	1,851,272	229,235	57,554,339	277
>10 rapid	660,540	1,321,080	133,661	38,580,214	277

Table 4: Escala de los artefactos de referencia estratificados después de la alineación con el Esquema de funciones de tiempo de ejecución común de 277 columnas. un "fila jugador-juego" significa un juego desde la perspectiva de un jugador, por lo que la mayoría de los juegos producir dos filas de juego de jugador.

Sólo para 3+0 blitz, el punto de referencia de producción actual utiliza 479.974 juegos. involucrando a 129,398 jugadores únicos y 28,138,904 filas distintas de estados de movimiento. el La tabla de características del juego de jugador relámpago alineada contiene 277 variables derivadas. el Los artefactos de creación rápida de características son aún más amplios en la etapa de estado de movimiento: el completo La tabla de estado de movimiento rápido utilizada durante la construcción de funciones contiene 168 columnas de estado de movimiento antes de la agregación en una función de juego de jugador de 277 columnas mesa. Los artefactos de referencia del tiempo de ejecución del frontend para los cuatro tipos de juegos son luego alineado con el mismo esquema de características del juego de jugador de 277 columnas, incluido el variables de desglose de taxonomía táctica ampliada. A partir de estas tablas de características, el El informe selecciona y muestra las métricas de coaching que son más interpretables.

Estos recuentos son importantes para interpretar los resultados. Las curvas de referencia No

son estimaciones de unos cientos de juegos revisados manualmente. En cambio, ellos provienen de decenas de millones de estados de la junta evaluados y más de cuatro millones de filas de juegos de jugadores en las cuatro categorías de juegos de producción actuales.

5 Descripción general de la tubería

El proyecto de construcción de referencia tiene seis pasos principales:

1. Seleccione juegos calificados Lichess en el tipo de juego objetivo y los estratos de calificación.
2. Analiza cada movimiento del juego, movimiento por movimiento, y almacena el movimiento anterior y posterior. estado del tablero.
3. Calcular variables intermedias de estado de movimiento desde el tablero, el movimiento, la perspectiva del jugador y la perspectiva del oponente.
4. Agregue variables de estado de movimiento en variables de características del juego del jugador.
5. Convierta las variables del juego del jugador en métricas de informes con una clara dirección de interpretación: más alto suele ser mejor, más bajo suele ser mejor o descriptivo.
6. Promedir cada métrica por cubo Elo e inspeccionar manualmente el resultado relación.

Esta estructura es importante porque muchas métricas de informes no están directamente visible en un encabezado PGN. Requieren reconstruir la posición, determinar qué piezas se movieron desde qué casillas iniciales, comparando el movimiento previo y el material y estructura después del movimiento, y luego agregar el resultado del mueve el nivel al nivel del juego del jugador.

6 Variables intermedias

Las variables intermedias están diseñadas para describir lo que cambió en el tablero. y por qué es importante desde la perspectiva del jugador del informe. Los ejemplos incluyen:

- número de movimiento, número de movimiento del jugador, lado a mover, color del jugador y jugador color;
- de escuadra, a escuadra, tipo de pieza móvil, tipo de pieza capturada y si el movimiento fue una captura, control, enroque, ascenso o mate;
- balance de materiales antes y después del movimiento;
- si una pieza menor abandonó su casilla inicial en un número de movimiento determinado;
- si la misma pieza se movió en movimientos consecutivos del jugador en el apertura;
- si la reina se movió antes del movimiento 10;
- lado del enroque, número de movimiento de enroque y si las torres se conectaron después del enroque;

- piezas sueltas, piezas colgantes, peones doblados, peones aislados, pasados peones y peones atrasados;
- estado de archivo de torre, archivos abiertos, archivos medio abiertos y seguridad del rey características;
- oportunidades tácticas, tácticas ejecutadas, tácticas de presión material, ganancias de seguimiento garantizadas y ganancias realizadas.

El conjunto intermedio exacto varía según la etapa de construcción. Las bases de datos de informes en tiempo de ejecución mantienen solo las columnas necesarias para generar informes rápidamente, mientras que la función completa Los artefactos incluyen tablas de estado de movimiento más amplias y parches de características utilizados para construir las filas finales de funciones del juego del jugador.

7 Ejemplos de métricas de principio de apertura

Los principios de apertura son ejemplos útiles porque son fáciles de entender para los usuarios. entender y fácil de conectar a los cálculos de estado de movimiento.

7.1 Minor Pieces Developed by Move 10

Esta métrica cuenta cuántos de los cuatro caballos y alfiles de un jugador quedan. sus casillas de origen en el décimo movimiento del jugador. El cálculo intermedio utiliza el color del motor, el tipo de pieza móvil, la casilla inicial y el movimiento del jugador número. Para las blancas, las casillas locales son b1, g1, c1 y f1. Para las negras, son b8, g8, c8 y f8. Para cada juego de jugador, Elo+Chess cuenta la casa distinta casillas entre caballos y alfiles que movieron el jugador mueve 10.

Esta es una métrica de que cuanto más alto suele ser mejor. No afirma que todo menor La pieza siempre debe moverse en el movimiento 10, pero entre los jugadores en desarrollo la experiencia empírica La relación es clara: los jugadores que alcanzan niveles de Elo más altos tienden a desarrollarse. más piezas menores antes.

7.2 Repeated Piece Moves in the Opening

Esta métrica cuenta los primeros movimientos adicionales realizados moviendo la misma pieza nuevamente. de desarrollar algo nuevo. El cálculo intermedio utiliza el anterior casilla de destino, la casilla de origen actual y el número de movimiento del jugador. si la misma pieza se movió en el movimiento anterior del jugador y se mueve nuevamente dentro del Los primeros 10 jugadores se mueven, el recuento de movimientos repetidos aumenta.

Esta es una métrica de que "menos es generalmente mejor". Mide los tiempos de apertura desperdiciados en un manera sencilla. Algunos movimientos repetidos están tácticamente justificados, pero a escala la La curva del cubo captura la tendencia general: jugadores más fuertes en el objetivo. Los niveles gastan menos movimientos iniciales moviendo la misma pieza repetidamente.

7.3 Queen Touched Before Move 10

Esta métrica cuenta los movimientos de la reina antes del movimiento 10 del jugador. El cálculo utiliza el tipo de pieza móvil y el número de movimiento del jugador. Si la pieza en movimiento es la reina y el número de movimiento del jugador es como máximo 10, la cuenta aumenta.

Esto también es más bajo, generalmente mejor para el rango objetivo Elo+Chess. Temprano Los movimientos de la reina pueden ganar concesiones materiales o forzarlas, pero muchos jugadores en desarrollo sacar a la reina demasiado pronto, perder tiempos por ataques a la reina y retrasar desarrollo de piezas menores y seguridad del rey.

7.4 Castling and King Safety

Las métricas de enroque utilizan banderas de enroque, lado de enroque, número de movimiento de enroque y el estado del tablero después del enroque. Los ejemplos incluyen si el jugador enrocó, si el jugador enrocó en el movimiento 10, si el jugador enrocó en el flanco de rey o flanco de dama, si las torres estaban conectadas después del enroque y si las filas de torres estaban abiertas o medio abiertas después del enroque.

No todas estas métricas se interpretan de la misma manera. Finalización del enroque y el enroque en el movimiento 10 es generalmente más alto es generalmente mejor en el rango objetivo. El enroque en el flanco de dama o en columnas abiertas puede depender más del contexto. Es por eso que Elo+Chess no se limita a codificar una regla moral; construye un cubo curva e inspecciona qué relaciones son lo suficientemente estables para el coaching.

8 Curvas del cucharón e inspección manual

Después de calcular las métricas del juego del jugador, Elo+Chess las agrupa por Lichess Elo cubo y calcula los promedios del cubo. El promedio del cubo es el resultado empírico bruto. punto de referencia para esa métrica. Luego se utiliza una línea de tendencia en el informe para Asigne calificaciones de estilo Elo a los valores del usuario.

Las curvas se inspeccionan manualmente antes de ser utilizadas. Una métrica es más útil cuando el cubo significa mostrar una relación coherente en todo el rango de calificación servido por el producto. Una métrica es menos útil cuando es ruidosa, plana, altamente no monótonos o impulsados por casos especiales que son difíciles de explicar a un usuario.

Este paso de inspección manual es especialmente importante por encima de la parte superior Elo+Chess. corte. Los jugadores más fuertes saben cuándo romper los principios de apertura y lo hacen. efectivamente. Por encima del rango de principiante a avanzado temprano, la relación La diferencia entre los hábitos simples y la calificación a menudo se aplana o cambia de forma. el producto Por lo tanto, enfatiza el rango donde las relaciones empíricas son fuertes. y pedagógicamente útil.

9 Ejemplo de curvas de apertura de rango completo

La figura 5 muestra cuatro curvas de principio de apertura de blitz 3+0. Los puntos grises muestran el rango de clasificación Lichess más amplio. Puntos azules y la tendencia azul La línea muestra la gama 600–1600 Lichess utilizada por los principales entrenadores Elo+Chess. puntos de referencia. La línea roja discontinua marca el límite superior.

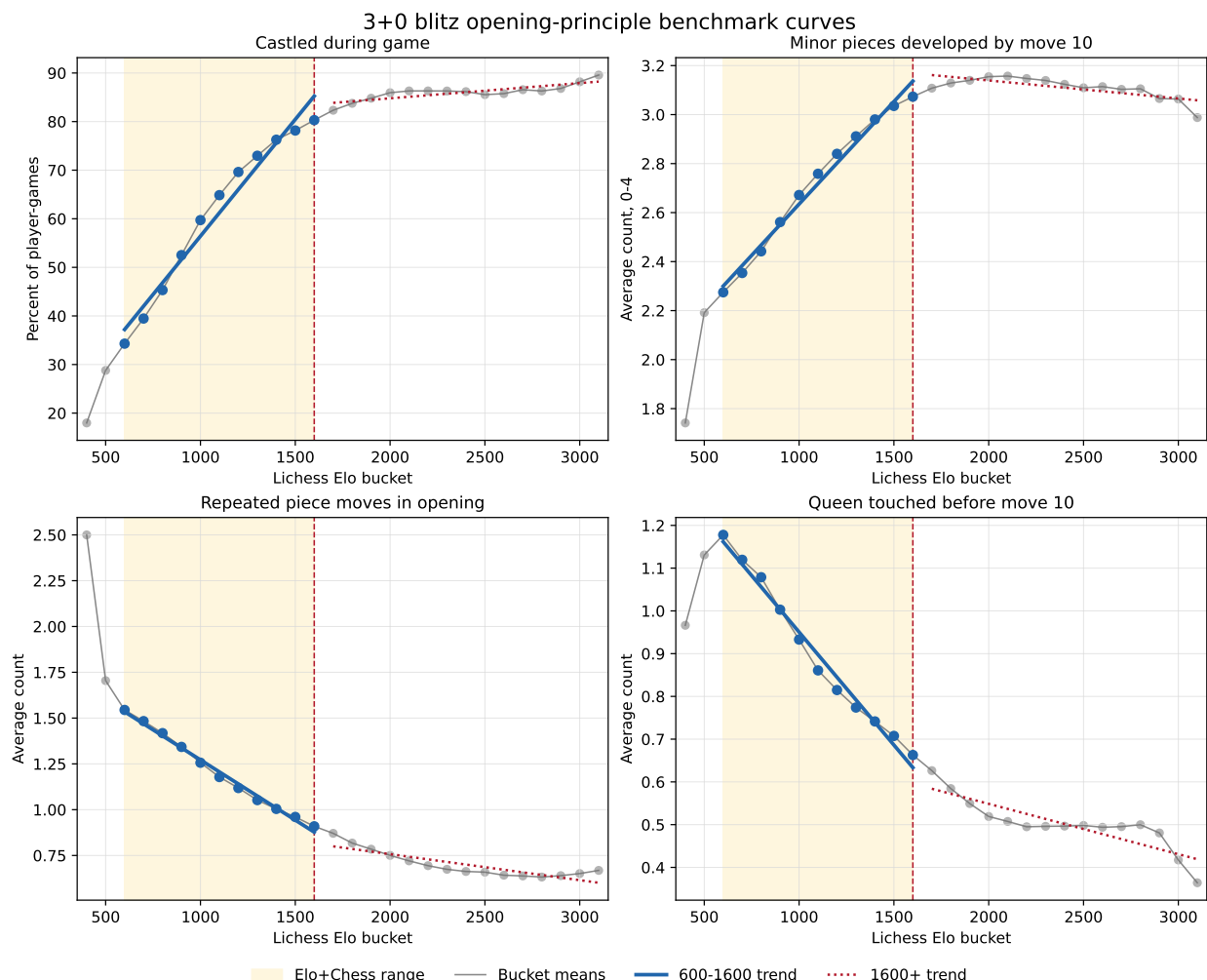


Figure 5: Curvas de referencia de principio de apertura seleccionadas para blitz 3+0. el Las relaciones son más sólidas y útiles en la gama de coaching Elo+Chess. Por encima del límite superior, varias relaciones se aplanan sustancialmente.

El mismo patrón es visible en el rango completo 10+0 rápido y en el más rápido. artefactos. La decisión central del modelado no ha cambiado: utilizar el método empíricamente rango estable para las afirmaciones de coaching y evite pretender que una simple curva de hábitos sigue siendo igualmente informativo para siempre.

El mismo patrón aparece numéricamente en la Tabla 5. Las pendientes son reportado como cambio de valor métrico por 100 puntos Elo. Enroque, pieza menor El desarrollo, el movimiento repetido de las piezas y el movimiento temprano de la reina muestran claramente movimiento a través de 600–1600. Por encima de 1600, las pendientes son mucho menores en estos ejemplos, lo cual es consistente con la idea de que los jugadores con calificaciones más altas rompen principios básicos de manera más selectiva y exitosa. Un segundo mecanismo es saturación. Para algunos hábitos, la métrica alcanza un techo práctico o dulce. lugar: los jugadores fuertes ya han aprendido el hábito, por lo que no hay razón para la curva siga moviéndose bruscamente hacia arriba. Más tampoco siempre es mejor después ese punto. Por lo tanto, los datos se aplanan porque los jugadores más fuertes pueden romper. reglas en posiciones concretas y porque mu-

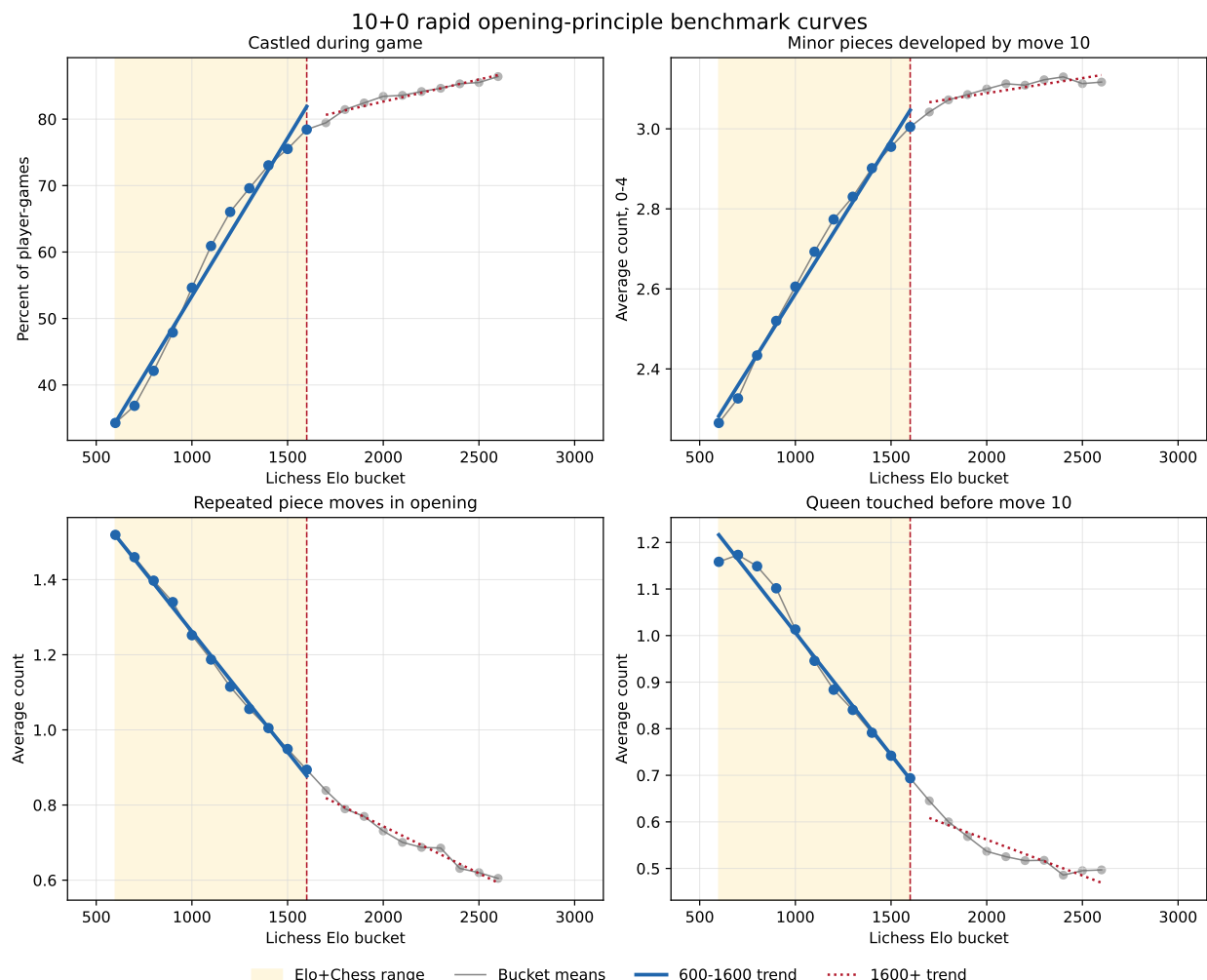


Figure 6: Curvas de referencia de principios de apertura seleccionadas para juegos rápidos 10+0. Las relaciones son más sólidas y útiles en la gama de coaching Elo+Chess. Por encima del límite superior, varias relaciones se aplanan sustancialmente.

Los hábitos básicos ya han alcanzado su rango operativo útil.

10 Por qué es importante el límite superior

El usuario objetivo de Elo+Chess es un jugador desde principiante hasta avanzado. Para estos jugadores, los hábitos básicos están fuertemente asociados con la calificación: desarrollar piezas, evitar movimientos tempranos repetidos de piezas, evitar movimientos tempranos innecesarios de la reina, enrocar tiempo, conectar torres, gestionar el material y evitar dejar piezas sueltas o colgando.

En calificaciones más altas, esos mismos hábitos superficiales se vuelven menos diagnósticos. un fuerte El jugador puede retrasar el enroque porque el centro está cerrado, mover la reina antes porque la posición contiene una táctica concreta, o mover la misma pieza dos veces porque gana un ritmo o fuerza una concesión estructural. El crudo hábito todavía tiene significado, pero la

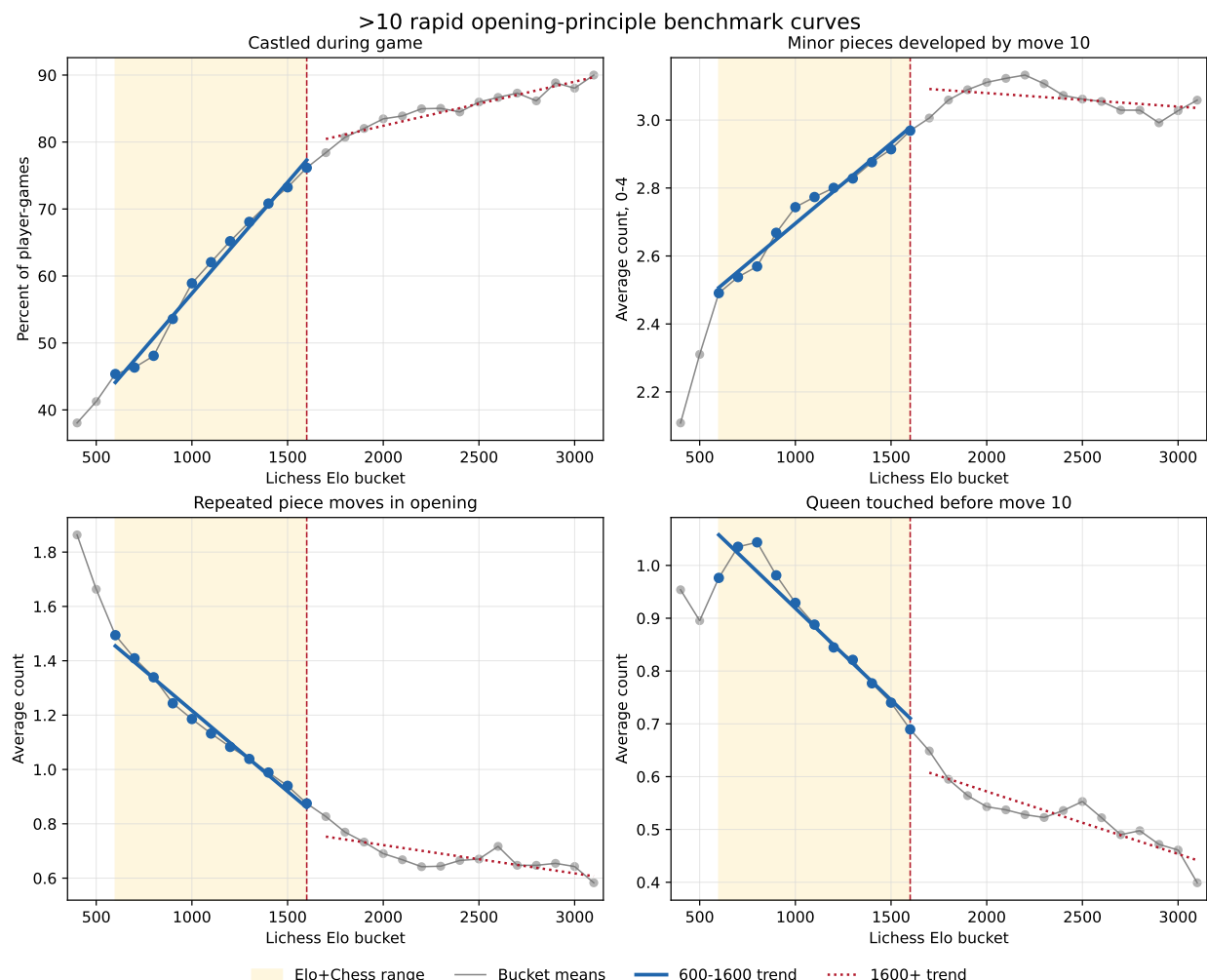


Figure 7: Curvas de referencia de principios de apertura seleccionadas para juegos rápidos más largos.

relación entre el hábito y la calificación se vuelve más condicional. Es por eso que Elo+Chess utiliza un límite superior para su entrenamiento principal. afirmaciones y por qué se advierte a los usuarios por encima de ese rango que la estrecha relación entre hábitos basados en principios y calificación se derrumba.

11 Métricas tácticas bilaterales y objetivos percentiles

Algunas métricas tienen la forma opuesta a la de los ejemplos del principio de apertura. Horquillas son un ejemplo útil porque el evento es inherentemente bilateral. Un jugador que encuentra y convierte bifurcaciones más útiles que sus pares está haciendo algo bueno, pero la tasa de población de las bifurcaciones convertidas cae con la calificación porque es más fuerte Los oponentes permiten menos oportunidades de bifurcación y se defienden contra amenazas tácticas. más activamente.

La figura 8 muestra este efecto. Ambas horquillas convertidas por el informe del jugador y

Game type	Metric	600–1600 slope per 100 Elo	1600+ slope per 100 Elo
1+0 bullet	Castled during game	4.137	0.550
	Minor pieces developed by move 10	0.063	0.009
	Repeated piece moves in opening	-0.051	-0.026
	Queen touched before move 10	-0.040	-0.011
3+0 blitz	Castled during game	4.804	0.313
	Minor pieces developed by move 10	0.084	-0.007
	Repeated piece moves in opening	-0.066	-0.014
	Queen touched before move 10	-0.053	-0.012
10+0 rapid	Castled during game	4.753	0.665
	Minor pieces developed by move 10	0.076	0.008
	Repeated piece moves in opening	-0.064	-0.025
	Queen touched before move 10	-0.052	-0.015
>10 minute rapid	Castled during game	3.320	0.659
	Minor pieces developed by move 10	0.047	-0.004
	Repeated piece moves in opening	-0.059	-0.010
	Queen touched before move 10	-0.035	-0.012

Table 5: Pendientes aproximadas de la curva del cucharón para los cuatro ejemplos de principio de apertura métricas en todos los tipos de juegos de producción.

los tenedores convertidos por parte del oponente se vuelven menos frecuentes a medida que Las calificaciones aumentan. Eso no significa que ejecutar bifurcaciones sea malo. Significa el evento crudo. La tasa está controlada en parte por la tasa de error del oponente y la textura del juego. Para métricas De esta manera, Elo+Chess utiliza percentiles de historial de pares dentro del propio cubo Elo del jugador en lugar de depender únicamente de una tendencia global de cubo medio.

Para construir estos objetivos percentiles, Elo+Chess toma muestras de jugadores por tipo de juego y Elo Bucket, luego utiliza juegos de clasificación recientes para cada jugador de la muestra. el El conjunto de objetivos actual utiliza 11 grupos de 600 a 1600, 200 jugadores por cubo y hasta 50 juegos de clasificación recientes por jugador para cada producción tipo de juego. Esto proporciona 2200 muestras de jugadores por tipo de juego y aproximadamente 110.000 filas de juegos de jugadores por tipo de juego para el percentil dentro del segmento distribuciones.

Game type	Elo buckets	Players per bucket	Games per player	Player samples
1+0 bullet	11	200	49–50	2200
3+0 blitz	11	200	27–50	2200
10+0 rapid	11	200	37–50	2200
>10 minute rapid	11	200	47–50	2200

Table 6: Conjuntos de objetivos de percentiles actuales del historial de pares. Estos son muestras de historial de jugadores utilizadas para estimar las distribuciones dentro del grupo para métricas de estilo percentil.

La figura 9 ilustra el objetivo resultante Distribución para la tasa de bifurcación convertida de jugadores muestreados en blitz 3+0. Cada caja es una distribución dentro del grupo entre los jugadores muestreados, no una media del grupo entre juegos individuales. En la interfaz de usuario del informe, el valor de un usuario se compara con el valor relevante. cuadro para el depósito Elo mapeado del usuario. Ésta es la base de afirmaciones como si la tasa de bifurcación convertida de un jugador es alta o baja en relación con sus pares en el mismo nivel, incluso cuando la tasa de eventos de la población general disminuye con la calificación.

Two-sided converted-fork rates by Elo bucket

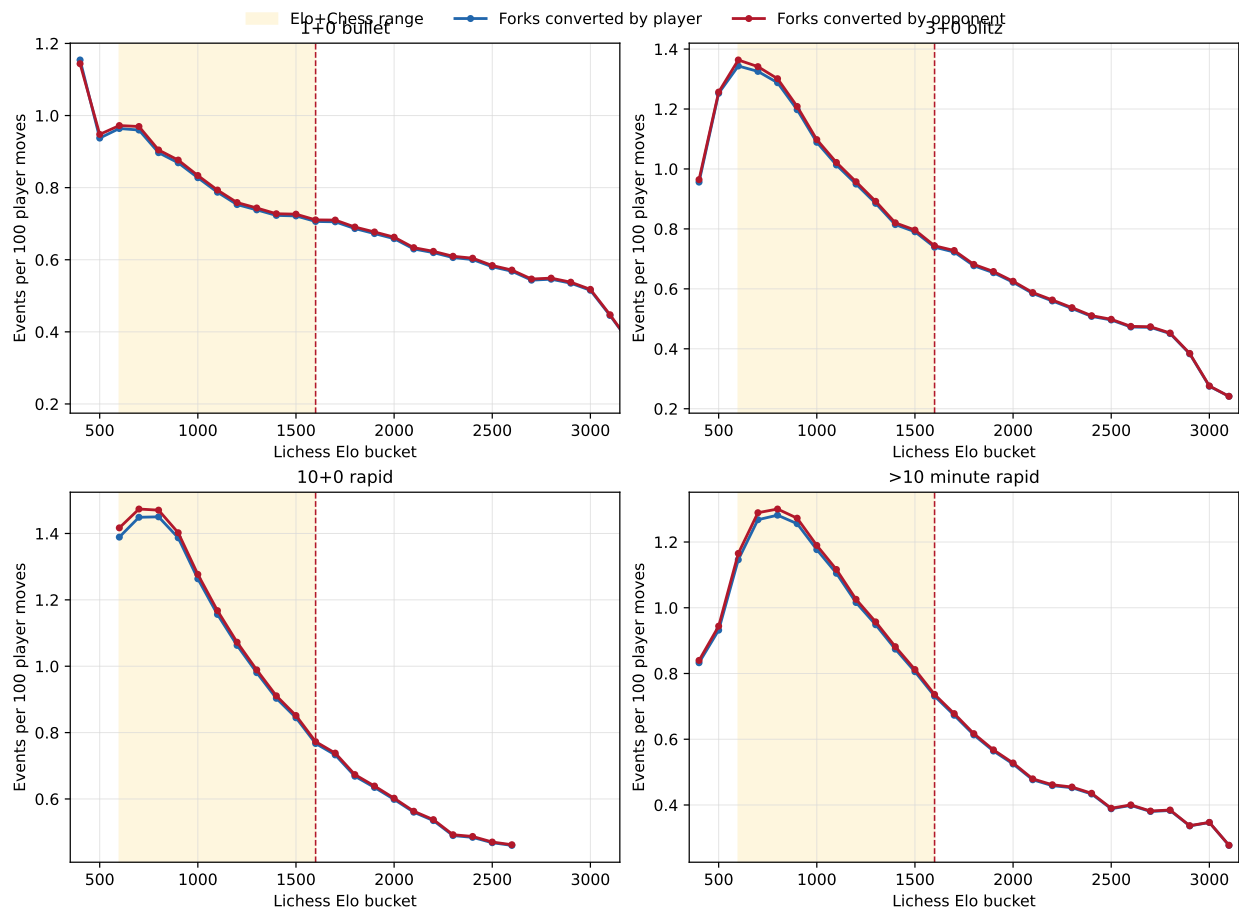


Figure 8: Horquillas convertidas de dos caras mediante cucharón Elo. Eventos de bifurcación convertidos caen con calificaciones más altas porque ambos lados defienden mejor y menos teneadores limpios conversiones están disponibles.

12 Costo de replicación

El método es reproducible en principio, pero no es ligero. Rebuilding las curvas de referencia requieren:

- acceso a grandes archivos mensuales de historial de juegos Lichess;
- filtrado estratificado por tipo de juego, grupo de clasificación y elegibilidad normas;
- un analizador de movimientos legales y un evaluador de junta-estado capaz de expandirse decenas de millones de estados en movimiento;
- característica lógica para apertura, material, estructura de peones, seguridad del rey, torre actividad, tácticas, gestión del tiempo y resultados del juego;
- agregación del estado de movimiento a características del juego del jugador;
- resúmenes a nivel de segmento para cada métrica del informe;

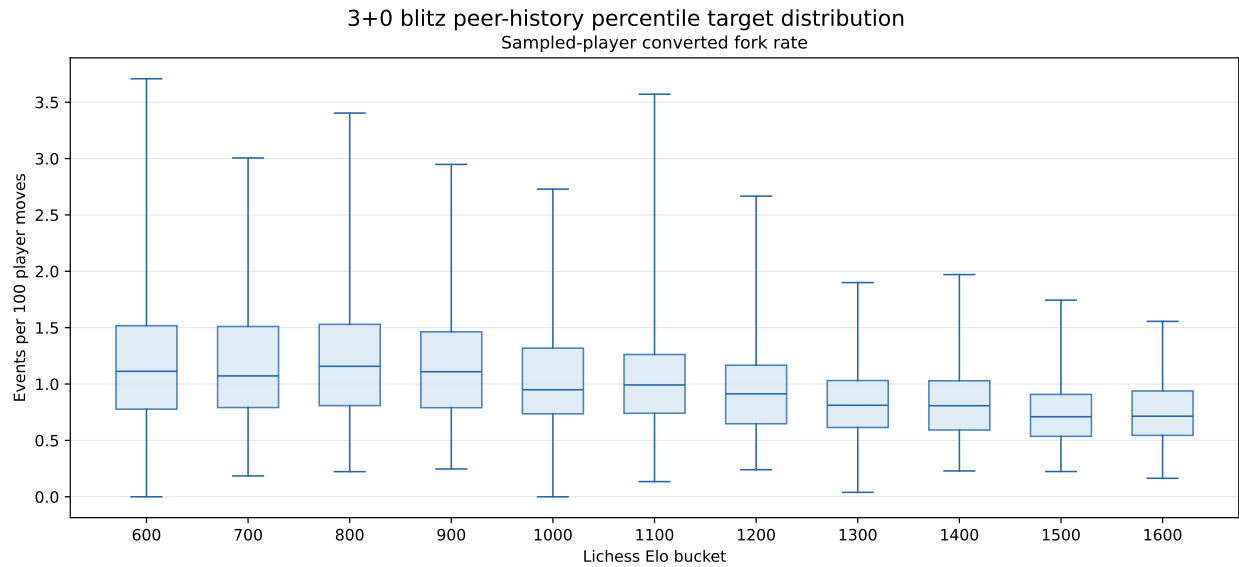


Figure 9: Ejemplo de distribución objetivo de historial de pares de blitz 3+0 para jugadores muestreados tasa de horquilla convertida. Los percentiles se calculan dentro del grupo Elo para que el informe compara a un jugador con compañeros que enfrentan oportunidades tácticas muy similares tarifas.

- revisión manual de cada curva de referencia candidata antes de utilizarla en un informe de entrenamiento.

Para blitz 3+0, esto significa casi medio millón de partidas, casi un millón filas de juegos de jugadores y más de 28 millones de filas de estados de movimiento. A través de los cuatro tipos de juegos de referencia, los artefactos de referencia actuales contienen más de 2,55 millones de juegos, más de 5,10 millones de filas de juegos de jugadores y más de 147 millones de filas de estado de movimiento. La escala es fundamental para el valor de la punto de referencia: permite que el informe muestre a los usuarios lo que realmente sucede en los juegos jugados por jugadores cercanos a su nivel en lugar de confiar únicamente en lemas generales del ajedrez.

13 Limitaciones

Las curvas de referencia son descripciones empíricas, no pruebas causales. un jugador no obtiene calificación simplemente haciendo coincidir una media de cubo. Las métricas también son Resúmenes imperfectos de posiciones complejas: un movimiento temprano de dama puede ser excelente. en una posición e imprudente en otra. El propósito de las curvas es identificar hábitos recurrentes y compararlos con el comportamiento de sus pares, no reemplazar el cálculo o el juicio de entrenamiento.

Las curvas también son curvas de escala Lichess. Cuando los usuarios traen el juego Chess.com historiales, Elo+Chess asigna el grupo de calificación relevante al punto de referencia Lichess escale utilizando el modelo de mapeo multiplataforma separado. Esa capa de mapeo es descrito en la nota de investigación complementaria Elo+Chess sobre Lichess-to-Chess.com conversión de calificación: [Estimating Cross-Platform Chess Rating Mappings with Regresión modal](#). Este mapeo no debe leerse como preservando el percentil. rango. Preserva una equivalencia de califi-

cación estimada entre las dos calificaciones. sistemas; El rango percentil sigue siendo específico de la plataforma porque el jugador activo las distribuciones son diferentes.

14 Gráficos del informe de producción

El informe de producción convierte los artefactos de referencia en paneles de nivel métrico que se puede leer sin conocer el proceso de compilación completo. Figuras 10 and 11 Muestre tres ejemplos del diseño del informe Elo+Chess en vivo.

Los paneles de principio de apertura en la Figura 10 utilice la presentación ordinaria de la curva de referencia. Cada tarjeta comienza con el familia de métricas y nombre de la métrica, luego proporciona una breve regla de interpretación. el El grado grande en la esquina superior derecha está separado intencionalmente del crudo. Valor métrico: responde a la pregunta del usuario "¿Cómo se clasifica este hábito?". ¿Contra el punto de referencia?", mientras que las filas siguientes conservan los valores medidos. Las cuatro filas de resumen muestran la duración, los últimos 10, las victorias y las derrotas. Esto hace que el Panel de diagnóstico en lugar de meramente decorativo. Si Lifetime y Last 10 están de acuerdo, el hábito probablemente sea estable. Si las ganancias y las pérdidas divergen, la métrica puede ser dependiente del resultado o vinculado al estado del juego en lugar de un simple hábito.

El área del gráfico utiliza un fondo oscuro restringido, una línea de tendencia dorada y colores marcadores para las cuatro rebanadas. La anotación de tendencia establece explícitamente si la relación de referencia sube o baja con Elo. Esto es importante porque Algunas métricas son "más alto es mejor", como el desarrollo de piezas menores, mientras que otros son más bajos, mejores, como el movimiento temprano de la reina. La tarjeta por lo tanto no le pide al usuario que deduzca la dirección únicamente a partir de la pendiente. la final Las líneas "Qué significa esto" y "Consejos de coaching" traducen el resultado de la evaluación comparativa en una instrucción de ajedrez concreta preservando al mismo tiempo el contexto empírico.

La figura 11 muestra el estilo percentil presentación utilizada para métricas cuya tasa bruta no está bien representada por un tendencia mundial única. Los tenedores ganadores son un ejemplo táctico: la tasa de población de los tenedores convertidos cambian tanto con la habilidad del jugador como con la tasa de error del oponente. el Por lo tanto, la tarjeta compara el valor del usuario con los historiales de jugadores pares en el cubo Elo mapeado. El diagrama de caja muestra la distribución por pares, incluida la rango intercuartílico, mediana y bigotes. El marcador del usuario se traza en el misma escala y etiquetados con valor bruto y rango percentil. adyacente Los depósitos de comparación permanecen visibles en forma silenciada para que el usuario pueda ver que el El grupo de pares es una escala local, no universal.

Este diseño hace que las métricas percentiles sean auditables. El usuario ve el valor bruto, el percentil, el grupo de pares utilizado para la comparación y el traducido Asignación de depósitos de Chess.com a Lichess. La misma duración, últimos 10, victorias y Las filas de pérdidas se conservan para que aún se puedan verificar las puntuaciones de percentiles tácticos. para efectos de actualidad y dependencia de resultados.

15 Conclusión

Los informes de referencia Elo+Chess se crean a partir de un gran cálculo de estado de movimiento sobre muestras estratificadas Lichess. Cada métrica comienza como un conjunto concreto de estados del tablero. y variables de estado de movimiento, se convierte en una característica del

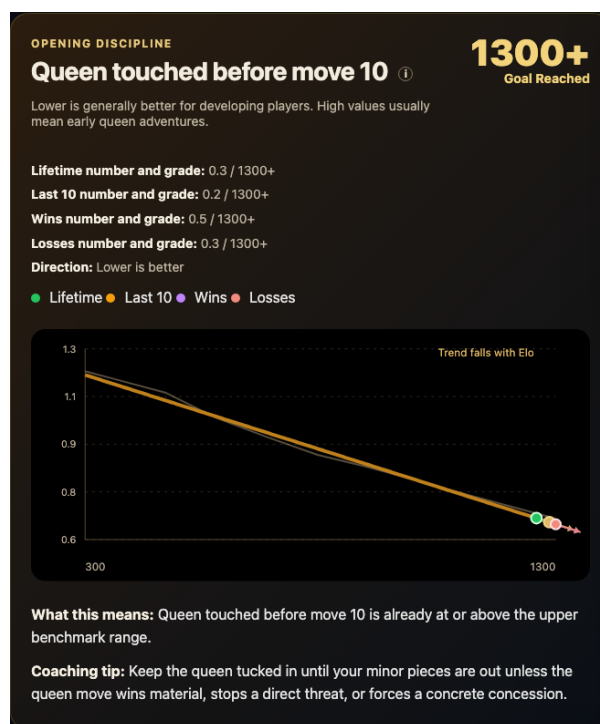
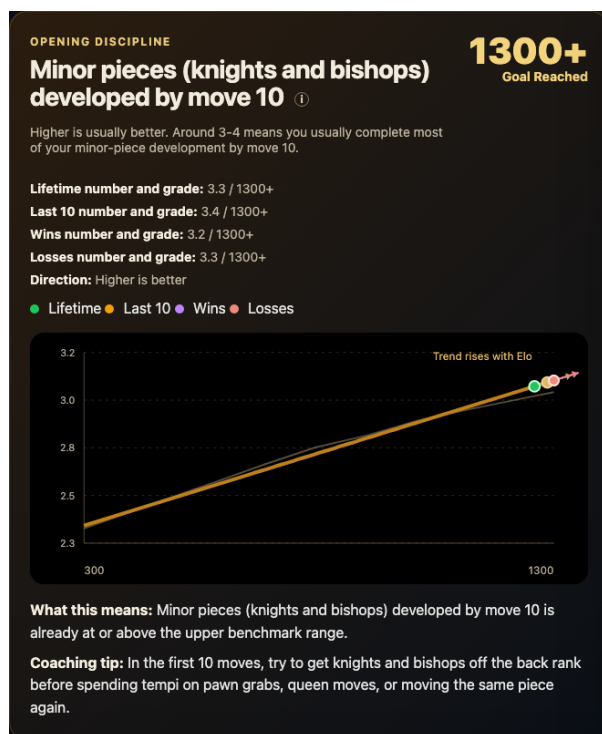


Figure 10: Paneles de principio de apertura de producción. La misma gramática visual maneja una métrica de más alto es mejor, desarrollo de piezas menores y una métrica de más bajo es mejor, movimiento temprano de la reina. Ambos paneles exponen la vida útil, los últimos 10, las victorias y las derrotas. para que el informe pueda separar los hábitos estables de los recientes o que dependen de los resultados. patrones.

juego del jugador y luego se promedia por cubo Elo para formar una curva de referencia empírica. El principio de apertura Los ejemplos muestran por qué este enfoque es útil: los hábitos ajedrecísticos simples tienen efectos fuertes y relaciones visibles con la calificación en el rango objetivo, mientras que aquellas las relaciones a menudo se aplanan o se vuelven más condicionales por encima del límite superior. Esa es la razón central por la que Elo+Chess centra sus afirmaciones de entrenamiento en la calificación. rango donde los datos los respaldan más claramente.

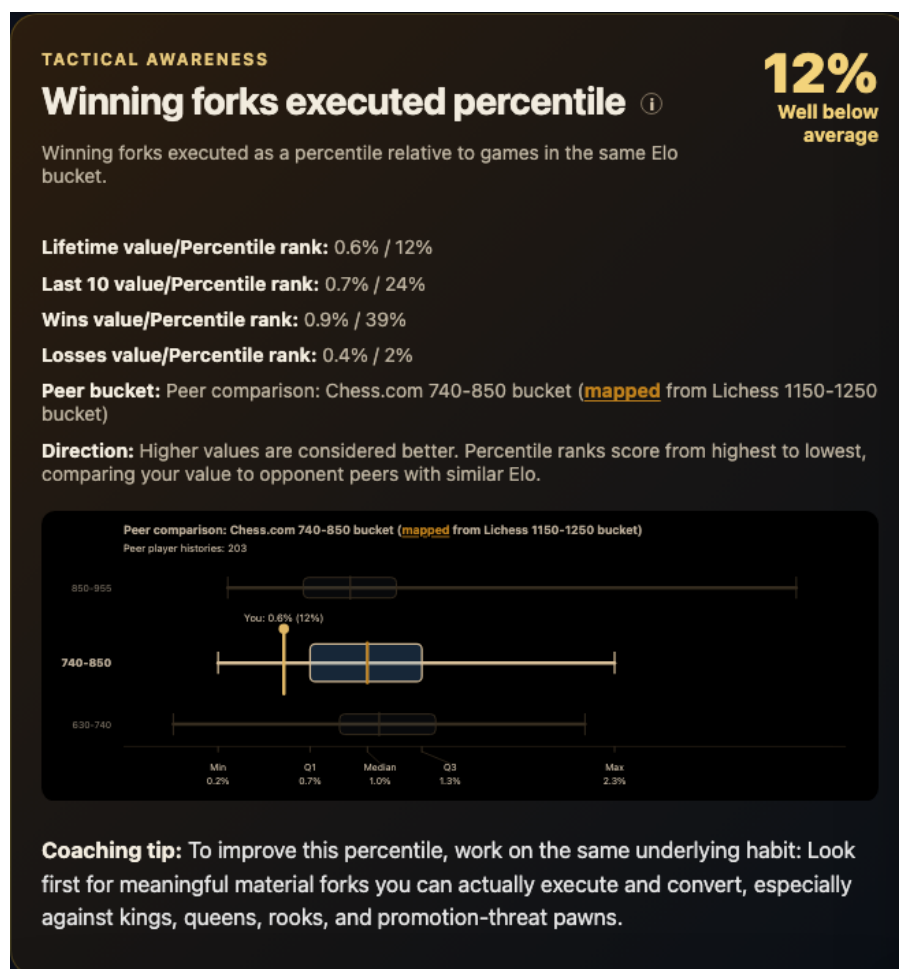


Figure 11: Panel percentil de producción de horquillas ganadoras ejecutado. El diagrama de caja muestra la distribución de pares del mismo segmento utilizada en el informe, mientras que el oro El marcador muestra el valor bruto del usuario y el rango percentil.