

Rétegzett Lichess benchmark görbék készítése a Elo+Chess számára

Elo+Chess Kutatási jegyzetek

2026. május 31

Abstract

A Elo+Chess benchmark jelentések nagy rétegzett mintákból épülnek fel Lichess.org játékok, amelyek a jegyzetben használt referenciamintát tartalmazzák 483 974 1+0 golyójáték, 479 974 3+0 villámjáték, 925 636 10+0 gyorsjáték, és 660 540 hosszabb rapid játék. Minden nagyobb játéktípus esetében kibővül a folyamat minden játék lépés előtti és lépés utáni tábla állapotba, több százat számol ki köztes és játékos-játék változók, ezeket a változókat összesíti coaching mérőszámokat, majd átlagolja ezeket a mutatókat a Lichess Elo csoport szerint. A minta szándékosan rétegzett, nem pedig arányos a nyers játékokkal pool: a cél a stabil viselkedési görbék becslése minden minősítési szinten, beleértve az egyébként ritkás vödröket is, hogy ne reprodukálják a természetet minősítési megoszlás. Az eredmény egy empirikus benchmark görbék családja: for minden mérőszámnál megkérdezhetjük, hogy a különböző Elo szinteken lévő játékosok milyen gyakran mutatnak egy adott értéket szokás vagy elv a valódi játékokban. Ez a jegyzet leírja a módszert, annak mértékét a számítás, és több nyitó elvi példák öntéssel, kisebb darabos fejlesztéssel, ismételt nyitóbábu mozgások és korai királynő mozgások. A hangsúly az módszertani: Elo+Chess nem feltételezi, hogy a sakk alapelvei univerzálisak lineáris minden készségszinten. Ehelyett a benchmark görbéket a besorolási tartomány, és a termék a kezdőtől a korai haladóiig terjedő régióra összpontosít ahol a kapcsolatok a leg-erősebbek és leginkább értelmezhetőek.

1 Cél

A benchmark rendszer célja, hogy a széleskörű sakktanácsokat átváltoztassa mérhető mennyiségek. Az olyan kijelentés, mint a „fejlessze a kisebb darabjait”. hasznos, de akcióbbá válik, ha meg tudjuk mondani, hány kisebb darab a A játékos általában a 10. lépéssel fejlődik, miként hasonlít ez a társaikhoz értékelési csoport, és hogyan változik ez a szokás az értékelési spektrumban.

A Elo+Chess ezért tényleges játékokból épít egy metrikakönyvtárat. Mindegyik mérőszám az minden játékos-játék sorra számítva, játéktípusonként és Elo vödör szerint összegezve, és benchmark görbéként használják a jelentésben. A jelentés ezután elhelyezheti a felhasználót ugyanazt a görbét, és megmutatja, hogy a felhasználó a viselkedés felett, alatta vagy közelében van-e hasonló játékosok körében figyelhető meg.

2 Forrásadatok és rétegződés

A benchmark forrás a nyilvános Lichess.org játékadatbázis. Elo+Chess használ nagy rétegzett minták játéktípus és értékelési csoport szerint, nem pedig kicsi kényelmi minta. A jelenlegi gyártási futásidejű melléktermékek négy fő területet fednek le játék kategóriák:

- 1+0 golyó,
- 3+0 blitz,
- 10+0 gyors,
- 10 percnél hosszabb gyorsjátékok.

A rétegződési cél a minősítési szintek közötti szélesség. Arányos véletlen minta hatékony lenne a teljes játékosállomány leírására, de igen nem lehet hatékony a benchmark görbék felépítésében. A leggyakoribb minősítési gyűjtők dominálna a mintában, míg az alacsony támasztékú farokvödörökben túl kevés lenne játékok a stabil metrikus átlagokért. A Elo+Chess ezért átlagos játék szerint vesz mintát Elo vödör. A jelenlegi tiszta minta konstrukcióban a nyers játékok állnak az első helyen szűrve érvényes minősítésű játékokra ismert játékosokkal, ismert értékelésekkel, érvényes eredményekkel, és ismert időszabályzók. A játékokat ezután a rendszer kiszűri az elegendő játékkal rendelkező játékosokra abban a sebességben és elég stabil értékelésben, hogy koherens képzettségi szintet képviseljen: minden játékos esetében a Elo+Chess kiszámítja a játékos sebességen belüli értékelési tartományát ($\max(\text{Elo}) - \min(\text{Elo})$) a forráskészleten keresztül, és megtartja játékok csak akkor, ha mindkét játékos rendelkezik legalább a beállított minimális játékszámmal és legfeljebb 300 Elo pont a sebességen belüli elsodródás. A drift szűrő nem arra való a normál javulás vagy eltérés megszüntetésére; csökkenti az eseteket, amikor a játékos a játékokban lényegesen eltérő képzettségi szintek keverednek, mint például az ideiglenes beszámolók, visszatérő játékosok, vagy olyan fiókok, amelyek gyorsan mozogtak több gyűjtőhelyen keresztül a mintavételi időszak. Végül a jogosult játékokat az egyes átlag-elokon belül rangsorolják kocka a játékosazonosító és a mag determinisztikus hash-ével, valamint a korlátozott számmal játékok megmaradnak minden egyes vödörből.

Ez a kialakítás sokkal hasznosabbá teszi a benchmark görbéket. Minden Elo vödör elegendő megfigyeléssel járul hozzá az egyes mutatók átlagos értékének becsléséhez, és a ritka, magas vagy alacsony besorolású vödöröket nem nyomja el a természetes forma az aktív játékos-eloszlásból. A Elo+Chess által használt benchmark görbék a következők majd a termék szempontjából leginkább releváns, nagyjából kezdő tartományra korlátozódik korai haladó játékon keresztül. A Lichess-skála jelentéseiben az az elsődleges tartomány 600-1600. Amikor a felhasználó a jelentést a Chess.com skálán tekinti meg, a a megfelelő minősítési csoportok ugyanarra a mögöttes Lichess-ra vannak leképezve benchmark csoportokat a külön platformok közötti értékelési leképezési módszerrel a kísérőpapírban leírtak szerint: [Estimating Cross-Platform Chess Rating Mappings with Modális regresszió](#).

3 Aktív játékosok értékelése

Ha megnézzük a Lichess.org lejáró Elo értékeléseinek megoszlását a végén 2026 márciusában a főbb játéktípusok szerint a következő számokat és százalékos értékeket látjuk. Ebben az elosztási nézetben minden játékost egyszer számol a rendszer pontos játéktípusonként, használja az adott játékos legutóbbi megfigyelt értékelését a megfelelő nyersben idővezérlő sorok, miután legalább öt játékot megköveteltek ebben a pontos játékban típus.¹

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 jelenítse meg a disztribúciókat ezzel a szűrővel és anélkül.

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: Legutóbbi megfigyelt aktív játékos Lichess értékelés pontos megoszlása játéktípust, miután legalább öt játékot megkövetel az adott játéktípusban.

Ezt az eloszlást nem szabad összetéveszteni a rétegzett referenciamintával. A disztribúció a raw-ban megfigyelt aktív Lichess játékosállományt írja le szkennelés. A benchmark mintát ezután szándékosan rétegezzük úgy, hogy az Elo vödröt elegendő támogatást nyújt a stabil metrikus görbékhez.

Egy adott játéktípusra vonatkozó Lichess.org Elo besorolásból egy egyenértékű Chess.com besorolás, a Elo+Chess a játéktípus-specifikus modált használja a társban becsült regressziós egyenletek [Elo+Chess rating-mapping paper](#):

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

Ezek az egyenletek értékelési értékeket képeznek, nem százalékos rangokat. Percentilis rangok meghatározott játékosállományon belül, és a Chess.com és Lichess játékos a medencék különböző formájúak. A 2 táblázatban a 10 perces gyorsjáték típus, az első oszlop egy kis mintát ad a lehetséges Chess.com Elo minősítési értékek. A második oszlop a megfelelőt adja meg A Chess.com százalékos rangsor a Chess.com-n látható 2026. május 31-én. A 71.7

a Lichess.org helyezés 10 perces rapidra a Lichess Elo besorolásnak felel meg 1,644², és ez az érték a harmadik oszlop. Ha alkalmazzuk a Chess.com–Lichess leképezési egyenletet, akkor 1,644 A Lichess körülbelül 1271-nek felel meg a Chess.com-n. Feltételezve a becsült A leképezési egyenlet ésszerű, a +522 különbség a Chess.com besorolás között az 1. oszlop és a 4. oszlop leképezett minősítése nagy számmal magyarázható alkalmi játékosok a Chess.com játékosállományban, felfűjják a Chess.com-t százalékos rang. Más szóval, a Lichess.org 72. percentilisben lenni nagyobb teljesítmény, mint a 72. percentilisben lenni a Chess.com-n, mert A százalékos rangok nem képezik le tisztán egyik rendszerről a másikra. Igazi Elo Az értékek leképezhetők az azonos játékos pontszámaiból becsült egyenletekkel a két rendszer.

A Table 2 hézagok oka az, hogy a leképezés egyenértékű játékerő-értékelést becsül, míg a százalékos leírja pozícióját egy platform saját lakosságán belül. Egy 71,7 százalékos játékos A Chess.com ezért nem automatikusan 71,7 százalékos játékos a Lichess-n, és fordítva. A megfigyelt számok összhangban vannak a Chess.com nyilvános adataival rapid pool sokkal nagyobb alacsonyabb besorolású vagy alkalmi játékos tömeggel, míg a aktív pontos vezérlésű Lichess pool koncentráltabb a tapasztalt ismétlés között játékosok. A kulcspon nem az, hogy egy százalékos skála "helyes" és a az másik "hibás"; ez az a százalékos rang nem hordozható platformok között kivéve, ha a mögöttes lejátszódisztribúciók azonos alakúak.

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: Szemléltető összehasonlítás a Chess.com 10 perces gyors percentilis között pont és ugyanez a százalékpont az aktív Lichess 10+0 rapidban elosztása legalább öt 10+0 játék megkövetelése után. “ Feltérképezett Chess.com ekv.” az aktuális Elo+Chess 10+0 gyors értékelési skála átalakítás azonos százalékos Lichess minősítésre. A legmagasabb farkú sorok szerepelnek a skála megkülönböztetésének megjelenítésére, de az illesztett minősítő vonal a legmegbízhatóbb a fő coaching tartományban.

Note on the 1500 spike. The visible jump near 1500 van jelen az adatokban, nem pedig simítással jön létre. A legújabbban megfigyelt értékelési eloszlás, pontosan 1500 értékelés az 1+0 13,1

bullet lejátszók, 5,6

17,6

kezdő vagy ideiglenes horgony a minősítési rendszerben.

A pontos 1500-as túske nagyrészt enyhén aktív fiókok műterméke. To ellenőrizze ezt, ugyanazokat a legfrissebb besorolású disztribúciókat számoltuk újra, miután megköveteltük minden játékosnak legalább öt játéka legyen a pontos játéktípusban. A túske majdnem eltűnik a szűrő alatt: pontosan 1500 értékelés esik 5,6-17,6

játékosok száma az összes játékosra kiterjedő elosztásban 0,2-1,2 megoszlása a 1 táblázatban szerepel.

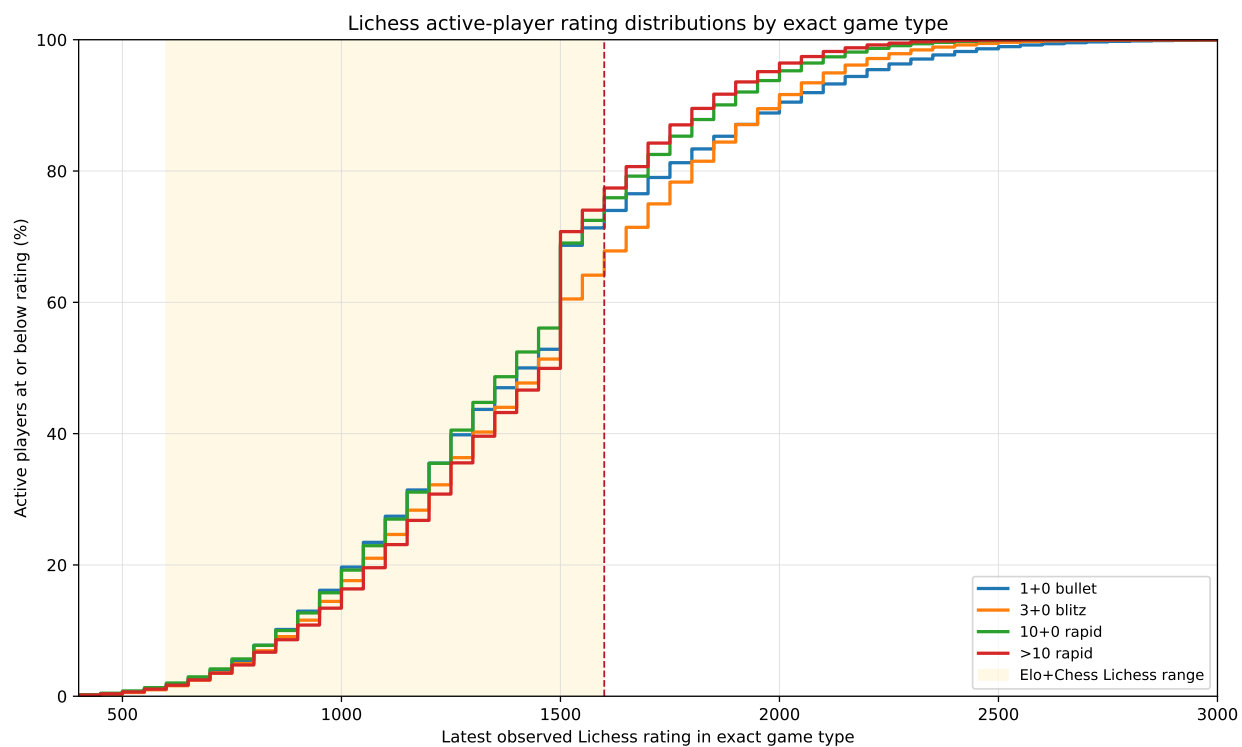


Figure 1: Az aktív játékosok összesített értékelései pontos játéktípusok szerint. A sárga sáv jelzi a fő Elo+Chess által használt 600–1600 Lichess sorozatot edzői benchmarkok.

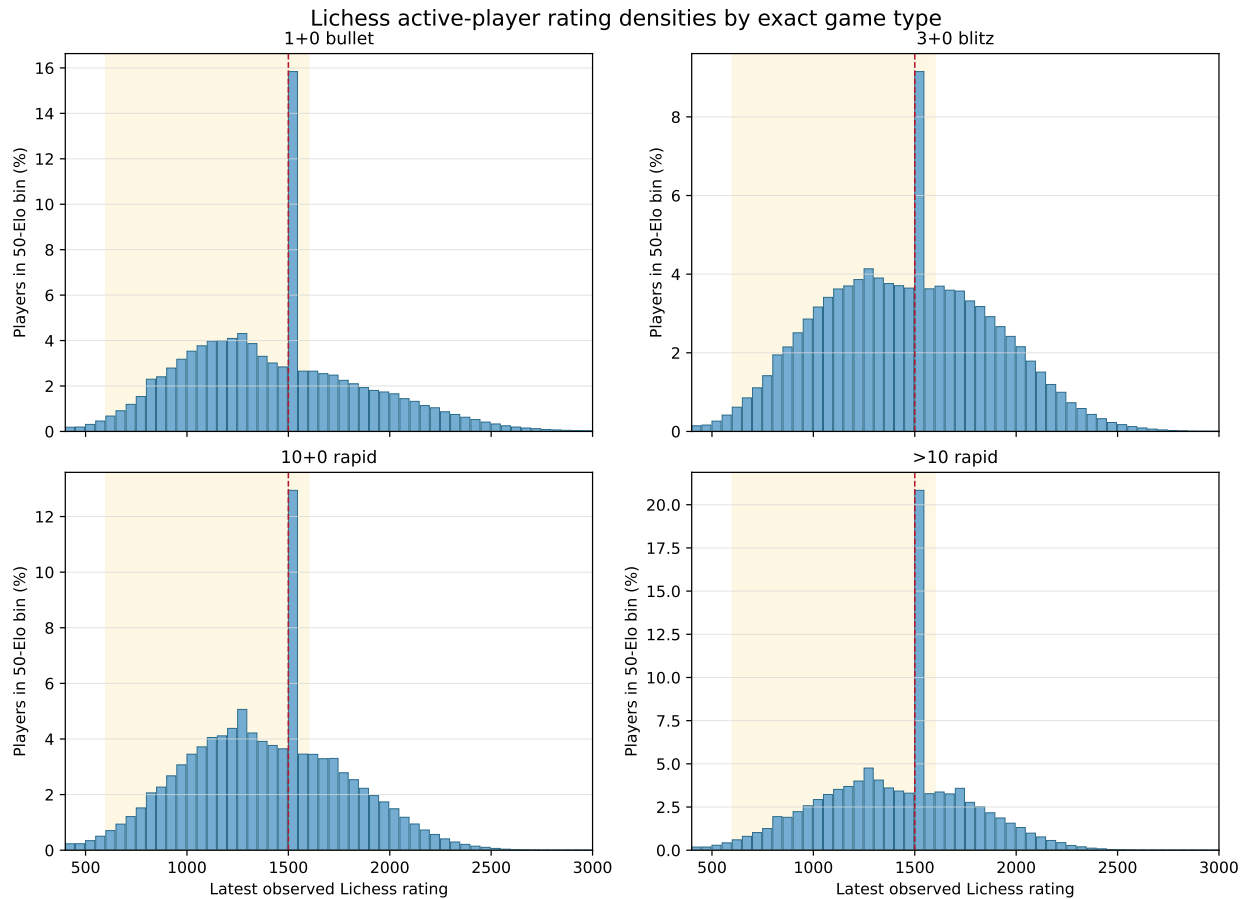


Figure 2: Azonos aktív játékosok értékelési eloszlásának sűrűsége. A szaggatott vonal jelzi az 1500 Lichess-t, ahol a nyers legfrissebb minősítésű disztribúciók egy nagy pontos névleges tömegpont. A sárga sáv a 600–1600 Lichess típust jelöli a fő Elo+Chess metrikajelentések által használt tartomány. A túske leginkább benne látható az összes játékos sűrűsége, mert sok enyhén aktív fiók pontosan a következőnél marad 1500; A 4 ábra megismétli a sűrűségdiagramot miután legalább öt játékot igényel a pontos játéktípusban, ami nagyrészt eltávolítja ezt a műtárgyat.

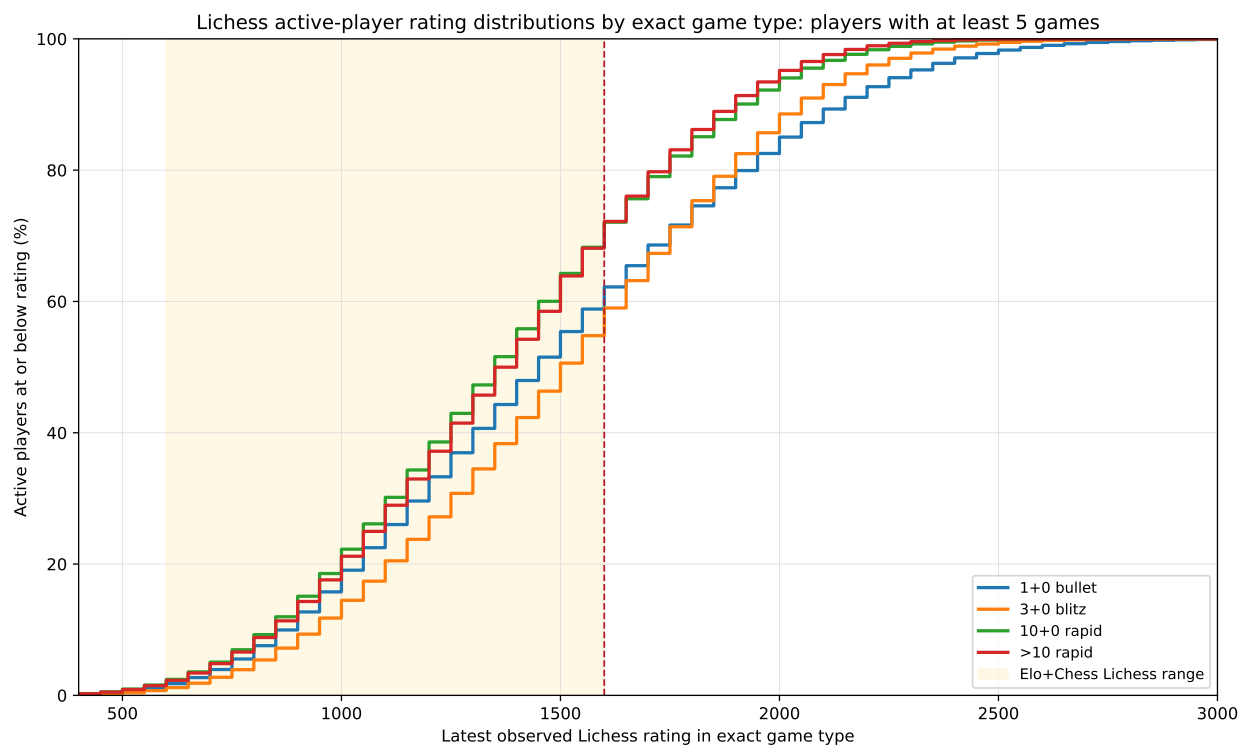


Figure 3: Aktív játékosok összesített értékelési eloszlása az at követelmény után legalább öt játék a pontos játéktípusban. Az 1500-as ugrás nagyon lecsökken.

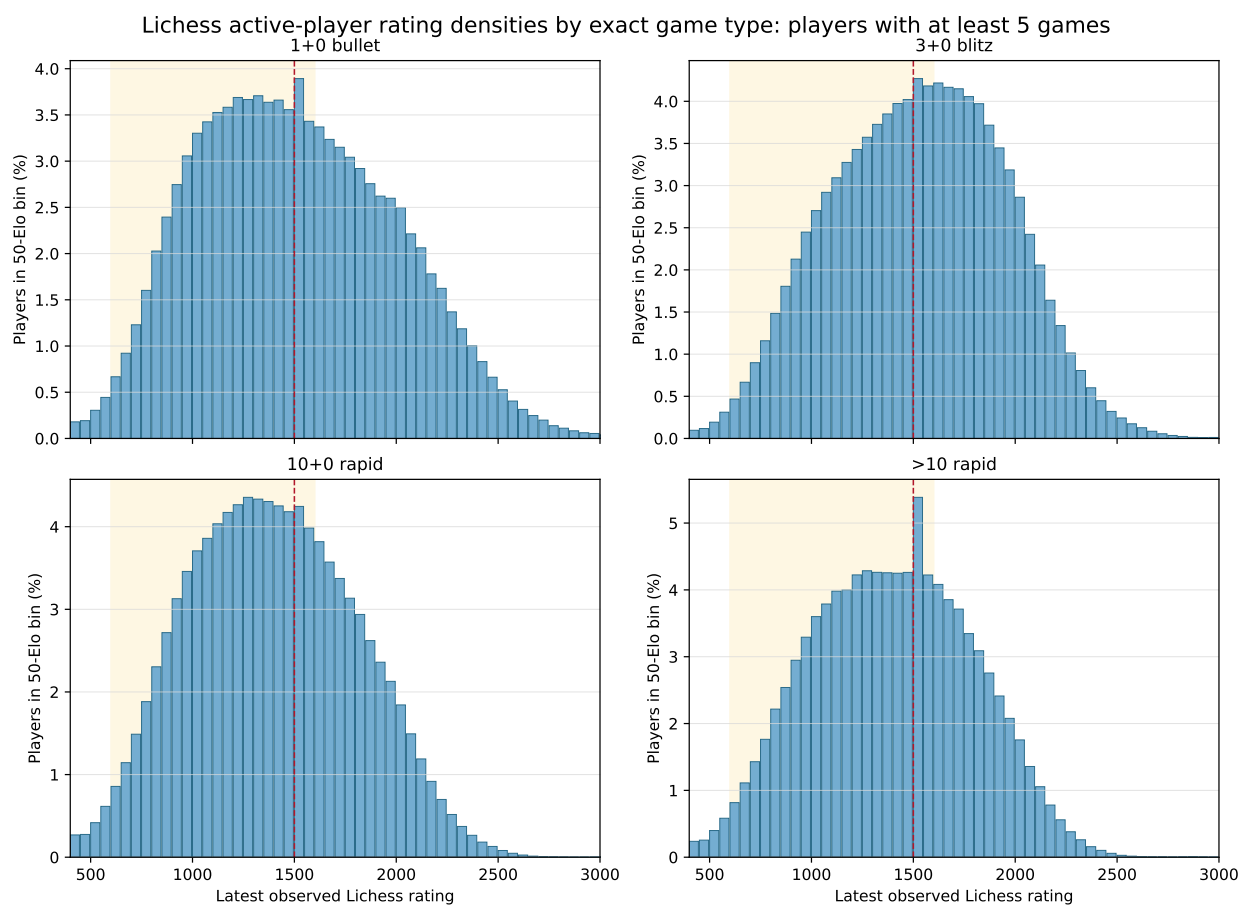


Figure 4: Sűrűségnézet, miután a pontos játékokban legalább öt játékra van szükség típus. Ez a nézet megerősíti, hogy a nagy 1500 tömegpont az összjátékosban a terjesztést többnyire az enyhén aktív fiókok hajtják. A sárga sáv jelzi a fő Elo+Chess metrikus jelentések által használt 600–1600 Lichess sorozat.

4 A jelenlegi benchmark műtermékek skálája

A számítás szándékosan nagy. Ez nem egy profilkaparás, egy kicsi letöltött PGN-készlet, vagy tanulságos példák kézzel válogatott gyűjteménye. A kezdeti Lichess bérlet teljes havi fájlokra vonatkozott 2025 januárjától 2026 márciusáig. A rétegezetlen keresési táblázat 1 382 178 087 játékot tartalmaz és 2 764 356 174 játékos-játék értékelési sor. Ezekben a sorokban 6 623 438 egyedi a játékosok felhasználói nevei legalább egyszer megjelennek.

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: Nem rétegzett Lichess szkennelési skála a benchmark mintavétel előtt. A szkennelés teljes havi játéktörténeti fájlokat takar 2025 januárjától márciusig 2026.

A gyártási benchmark műtermékek sokkal kisebbek, mint a nyers szkennelés, mert metrikus konstrukcióhoz választják ki. Még mindig nagyok: a viszonyítási alap színpad egy mozgatási állapotú bővítési és összesítési folyamat több százon játéktípusonként több ezer játék.

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: A rétegzett benchmark műtermékek skálája a következőhöz igazítás után közös 277 oszlopos futásidejű szolgáltatásséma. A „játékos-játék sor” egy játékot jelent egy játékos szemszögéből, tehát a legtöbb játék hozzon létre két játékos-játék sort.

Csak 3+0 blitz esetén a jelenlegi termelési referenciaérték 479 974 játékot használ 129 398 egyedi játékost és 28 138 904 különböző mozgásállapot-sort foglal magában. Az igazított blitz játékos-játék jellemző táblázat 277 származtatott változót tartalmaz. A gyors funkcióépítésű műtermékek még szélesebbek a mozgatási állapot szakaszában: a teljes A funkcióépítés során használt gyorsmozgás-állapottáblázat 168-at tartalmaz mozgat-state oszlopokat, mielőtt egy 277 oszlopos játékos-játék funkcióvá egyesítené táblázat. A frontend futásidejű benchmark műtermékei mind a négy játéktípushoz majd ugyanahhoz a 277 oszlopos játékos-játék funkciószámához igazítva, beleértve a kiterjesztett taktika-taxonómia kitörési változókat. Ezekből a jellemzőtáblázatokból a jelentés kiválasztja és megjeleníti a leginkább értelmezhető coaching mérőszámokat.

Ezek a számok fontosak az eredmények értelmezéséhez. A benchmark görbék nem néhány száz manuálisan felülvizsgált játék becslése. Ehelyett ők több tízmillió értékelt testületi államból és több mint négyből származnak millió játékos-játéksor a jelenlegi négy gyártási játékkategóriában.

5 Csővezeték áttekintése

A benchmark építési folyamat hat fő lépésből áll:

1. Válassza ki a besorolt Lichess játékokat a céljátéktípusból és a besorolási rétegből.
2. Elemezze az egyes játékokat lépésenként, és tárolja a lépés előtti és utáni lépést tábla állapotban.
3. Köztes mozgásállapot-változók kiszámítása a tábláról, a mozgás, a játékos perspektívája és az ellenfél perspektívája.
4. A mozgásállapot-változók összesítése játékos-játék jellemzői változókká.
5. Konvertálja a játékos-játék-változókat jelentési mérőszámokká, egyértelműen értelmezési irány: magasabb általában jobb, alacsonyabb általában jobb, vagy leíró jellegű.
6. Átlagolja az egyes mérőszámokat Elo-gyűjtőrészszel, és manuálisan ellenőrizze az eredményt kapcsolatban.

Ez a struktúra azért fontos, mert sok jelentési mérőszám nem közvetlenül látható a PGN fájlban. Megkövetelik a helyzet rekonstrukcióját, elhatárolását melyik bábu melyik kezdőnégyzetből mozdult el, összehasonlítva a mozgás előtti és utómozgatás anyag és szerkezet, majd az eredmény összesítése a lépés szintet a játékos-játék szintre.

6 Köztes változók

A köztes változók célja, hogy leírják, mi változott a táblán és miért számít ez a riportjátékos szempontból. Példák:

- lépésszám, játékos lépésszám, mozgatandó oldal, mozgató színe és játékos szín;
- négyzetből, négyzetbe, mozgó darab típus, rögzített darab típus és hogy a lépés elfogás, csekk, vár, előléptetés vagy társ volt-e;
- költözés előtti és költözés utáni anyagmérleg;
- egy mellékdarab elhagyta-e a kezdő mezőjét egy adott lépésszámmal;
- hogy ugyanaz a bábu mozgott-e az egymást követő játékosok lépései során a nyílás;
- hogy a királynő mozdult-e a 10. lépés előtt;
- vároldarab, a vármozgás száma, és hogy a bástya összekapcsolódott-e öntés után;
- laza darabok, függő darabok, megkettőzött gyalogok, elszigetelt gyalogok, átadva gyalogok és hátrafelé álló gyalogok;
- rook fájl állapota, nyitott fájlok, félig nyitott fájlok és king-safety jellemzők;
- taktikai lehetőségek, végrehajtott taktika, anyagnyomás taktika, garantált nyomon követési nyereség és realizált nyereség.

A pontos köztes készlet építési szakaszonként változik. Futásidejű jelentés adatbázisok tárolják csak az oszlopok szükségesek a jelentések gyors kiszolgálásához, míg a teljes szolgáltatásépítés A műtermékek közé tartoznak a szélesebb mozgatási állapotú táblák és a konstrukcióhoz használt jellemzőfoltok az utolsó játékos-játék jellemzősorok.

7 Példák a nyitóelv metrikájára

A nyitási alapelvek hasznos példák, mert könnyen használhatók a felhasználók számára érthető és könnyen csatlakoztatható a mozgásállapot-számításokhoz.

7.1 Minor Pieces Developed by Move 10

Ez a mérőszám azt számolja meg, hogy egy játékos négy lovagja és püspöke közül hány maradt el a játékos 10. lépésével dobja be a saját mezőit. A közbenső számítás használja a mozgató színét, a mozgó darab típusát, a kezdő négyzetet és a játékos mozgását számot. Fehér esetében a kezdőnégyzetek: b1, g1, c1 és f1. Fekete számára azok b8, g8, c8 és f8. Minden játékos-játéknál a Elo+Chess a különálló otthont számolja lovagok és püspökök közötti mezők, amelyeket a játékos lépései mozgattak 10.

Ez egy magasabb általában jobb mutató. Nem állítja, hogy minden kiskorú a bábunak mindig 10 lépésenként kell mozognia, de a fejlődő játékosok között az empirikus A kapcsolat egyértelmű: azok a játékosok, akik magasabb Elo vödröt érnek el, általában fejlődnek több kisebb darab korábban.

7.2 Repeated Piece Moves in the Opening

Ez a mérőszám azokat az extra korai mozdulatokat számolja, amelyeket ugyanazon bábu ismételt mozgatására fordítottak valami újat kifejleszteni. A közbenső számítás az előzőt használja cél mező, az aktuális kiindulási négyzet és a játékos lépésszáma. Ha ugyanaz a bábu az előző játékos lépésénél mozgott, és ismét a játékoson belül mozog Az első 10 játékos lépése esetén az ismételt lépések száma növekszik.

Ez egy alacsonyabb általában jobb mutató. Elpazarolt nyitási tempót mér a egyszerű módon. Néhány ismételt mozdulat taktikailag indokolt, de a léptékben a A vödörgörbe az általános tendenciát ragadja meg: erősebb játékosok a célpontban szintek kevesebb korai lépést töltenek ugyanazon bábu ismételt mozgatásával.

7.3 Queen Touched Before Move 10

Ez a mérőszám a dáma lépéseit számolja a játékos 10 lépése előtt. A közbenső A számítás a mozgó darab típusát és a játékos lépésszámát használja. Ha a mozgó darab a királynő és a játékos lépésszáma legfeljebb 10, a szám növekszik.

Ez is alacsonyabb-általában jobb a Elo+Chess céltartományban. Korán A királynő lépései anyagilag nyerhetnek vagy engedményeket kényszeríthetnek, de sok fejlődő játékos hozd ki a királynőt túl korán, veszítsd el a tempót a királynő elleni támadásoktól, és késlekedj kisebb darabos fejlesztés és királybiztonság.

7.4 Castling and King Safety

Az öntési mérőszámok öntési zászlókat, öntési oldalt, öntés lépésszámot és a tábla állapota öntés után. Példák arra, hogy a játékos várt-e, hogy a játékos a 10. lépéssel varakozott-e, a játékos a királyoldalt, vagy queenside, hogy a bástya össze volt-e kapcsolva az öntés után, és hogy a bástya akták az öntés után nyitva vagy félig nyitva voltak.

Ezeket a mutatókat nem egyformán értelmezik. A várkastély befejezése és a 10. lépéssel történő öntés általában magasabb-általában jobb a céltartományban. A Queenside casting vagy

a nyitott fájlokba öntés inkább kontextusfüggő lehet. Ezért van az, hogy a Elo+Chess nem egyszerűen kódol egy erkölcsi szabályt; vödröt épít görbék, és megvizsgálja, hogy mely kapcsolatok elég stabilak a coachinghoz.

8 Vödör görbék és kézi ellenőrzés

A játékos-játék mutatóinak kiszámítása után a Elo+Chess a Lichess Elo szerint csoportosítja őket vödör és kiszámítja a vödör átlagait. A vödör átlaga a nyers empirikus viszonyítási pont ehhez a mutatóhoz. Ezután egy trendvonalat használnak a jelentésben Elo-stílusú osztályzatokat rendelhet a felhasználói értékekhez.

Használat előtt a görbéket manuálisan ellenőrzik. A mérőszám hasznosabb amikor a vödör eszközök koherens kapcsolatot mutatnak az értékelési tartományban a termék szolgálja ki. A mérőszám kevésbé hasznos, ha zajos, lapos vagy erős nem monoton, vagy olyan speciális esetek vezérlik, amelyeket nehéz elmagyarázni a felhasználónak.

Ez a kézi ellenőrzési lépés különösen fontos a Elo+Chess felső rész felett levágás. Az erősebb játékosok tudják, mikor kell megszegni a nyitóelveket, és meg is teszik hatékonyan. A kezdőtől a korai haladóig terjedő tartomány felett a kapcsolat az egyszerű szokások és az értékelés között gyakran ellaposodik vagy alakot változtat. A termék ezért hangsúlyozza azt a tartományt, ahol az empirikus kapcsolatok erősek és pedagógiaiailag hasznos.

9 Példa teljes tartományú nyitási görbére

A 5 ábra négy 3+0 blitz nyitási elvű görbét mutat. A szürke pontok a szélesebb Lichess minősítési tartományt mutatják. Kék pontok és a kék trend vonal mutatja a fő Elo+Chess coaching által használt 600–1600 Lichess sorozatot benchmarkok. A szaggatott piros vonal a felső határt jelöli.

Ugyanez a minta látható a teljes tartományú 10+0 rapid és hosszabb-rapid esetében is műtárgyak. A központi modellezési döntés változatlan: használja az empirikus stabil tartomány a coaching állításokhoz, és ne tegyünk úgy, mintha egy egyszerű szokásgörbe lenne örökké ugyanolyan informatív marad.

Ugyanez a minta numerikusan jelenik meg a Table 5 táblázatban. Lejtők vannak 100 Elo-pontonkénti metrika-érték változásként jelentették. Castling, mellékdarab a fejlődés, az ismételt darabmozgás és a korai királynő mozgás mind tisztán mutat mozgás 600–1600 között. 1600 felett ezekben jóval kisebbek a lejtők példák, ami összhangban van azzal az elképzeléssel, hogy a magasabb minősítésű játékosok törnek alapelvek szelektívebben és sikeresebben. A második mechanizmus az telítettség. Egyes szokások esetében a mérőszám eléri a praktikus mennyezetet vagy édességet spot: az erős játékosok már megtanulták a szokást, így nincs miért a görbe élesen felfelé haladjon. A több sem mindig jobb utána azt a pontot. Az adatok ezért mindkettőt ellaposítják, mert az erősebb játékosok törhetnek szabályokat a konkrét pozíciókban és mert sok alapvető szokás már elérte hasznos működési tartományuk.

10 Miért számít a felső határvonal?

A Elo+Chess célfelhasználó kezdőtől korai haladóig terjedő játékos. Ezekre játékosok, az alapvető szokások erősen összefüggenek a minősítéssel: darabok fejlesztése, kerülje az ismételt korai

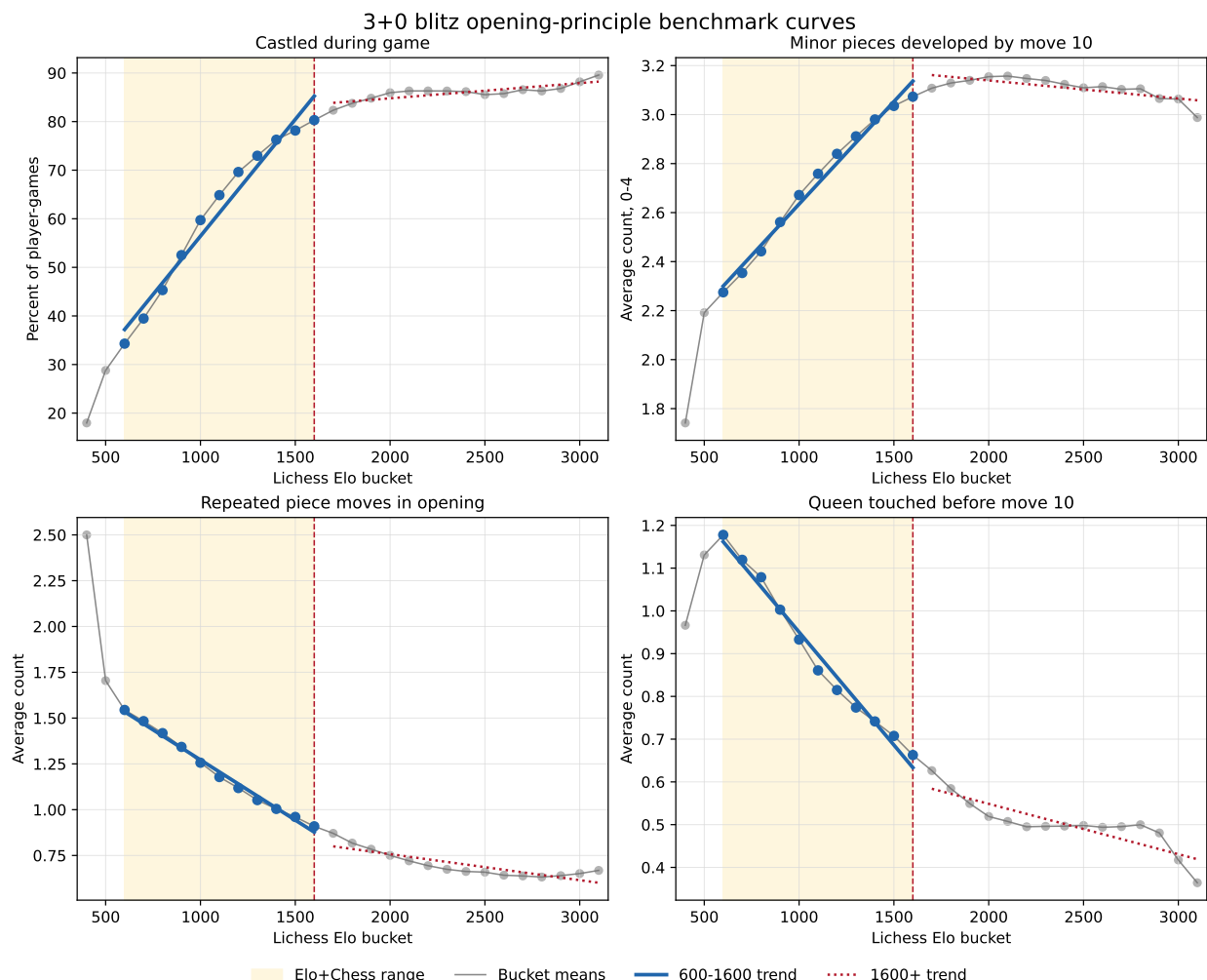


Figure 5: Válogatott nyitási elvű benchmark görbék 3+0 blitzhez. A kapcsolatok a legerősebbek és a leghasznosabbak a Elo+Chess coaching kínálatában. A felső határ felett számos kapcsolat lényegesen ellaposodik.

bábumozgásokat, kerülje a szükségtelen korai dámalépéseket, várat időben kösse össze a bátyákat, kezelje az anyagot, és ne hagyja lazán a darabokat lógó.

Magasabb értékeléseknél ugyanezek a felületi szokások kevésbé diagnosztikusak. Egy erős játékos késleltetheti a castingot, mert a központ zárva van, mozgassa a királynőt korán mert a pozíció konkrét taktikát tartalmaz, vagy mozgassd kétszer ugyanazt a darabot mert tempót nyer vagy szerkezeti engedményt kényszerít ki. A nyers szokás még mindig van értelme, de a szokás és az értékelés közötti kapcsolat egyre inkább erősödik feltételes. Ez az oka annak, hogy a Elo+Chess felső határértéket használ a fő edzéshez állítások, és miért figyelmeztetik a felhasználókat ezen tartomány felett, hogy a szoros kapcsolat elvi szokások és a minősítés között megtörik.

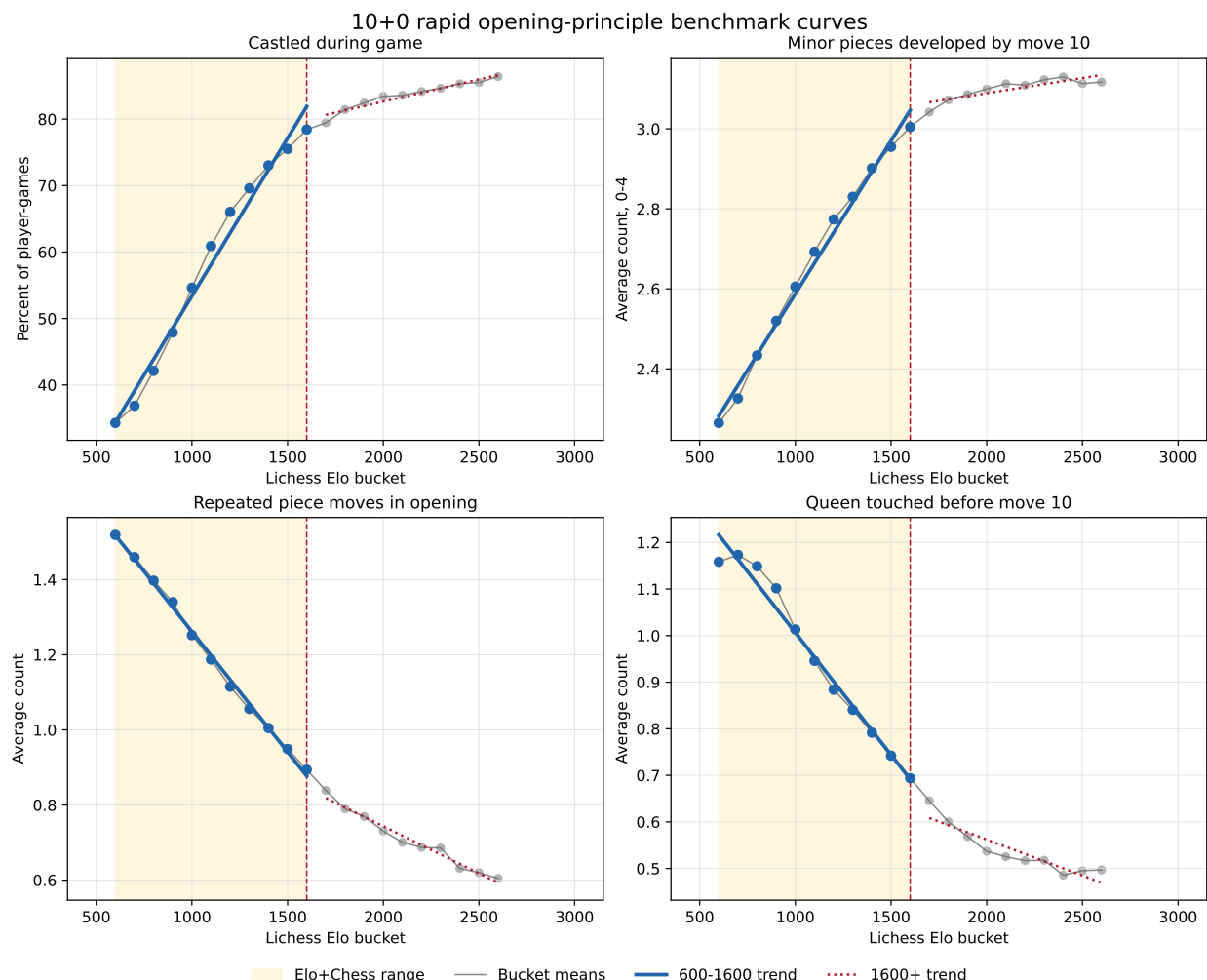


Figure 6: Válogatott nyitási elvű benchmark görbék 10+0 gyorsjátékokhoz. A kapcsolatok a leg-erősebbek és a leghasznosabbak a Elo+Chess coaching kínálatában. A felső határ felett számos kapcsolat lényegesen ellaposodik.

11 Kétoldalú taktikai mutatók és százalékos célok

Egyes metrikák a nyitóelv példáival ellentétes alakúak. Villák hasznos példa, mert az esemény eredendően kétoldalú. Egy játékos, aki hasznosabb villát talál és alakít át, mint a társak, valami jót tesz, de az átalakított villák népességi aránya a minősítéssel esik, mert erősebb az ellenfelek kevesebb villalehetőséget engednek meg és védekeznek a taktikai fenyegetésekkel szemben aktívabban.

A 8 ábra ezt a hatást mutatja. Mindkettő villát alakított át a riportjátékos és az ellenfél által konvertált villák ritkábbak lesznek, mint az értékelések emelkednek. Ez nem jelenti azt, hogy a villák végrehajtása rossz. A nyers eseményt jelenti az arányt részben az ellenfél hibaaaránya és a játék textúrája szabályozza. A mérőszámokhoz így a Elo+Chess ezért peer-history percentiliseket használ a a játékos saját Elo-vödőre, ahelyett, hogy csak a globális kocka-átlagos trendre hagyatkozna.

A százalékos célok felépítéséhez a Elo+Chess játéktípus és játéktípus szerint mintát vesz

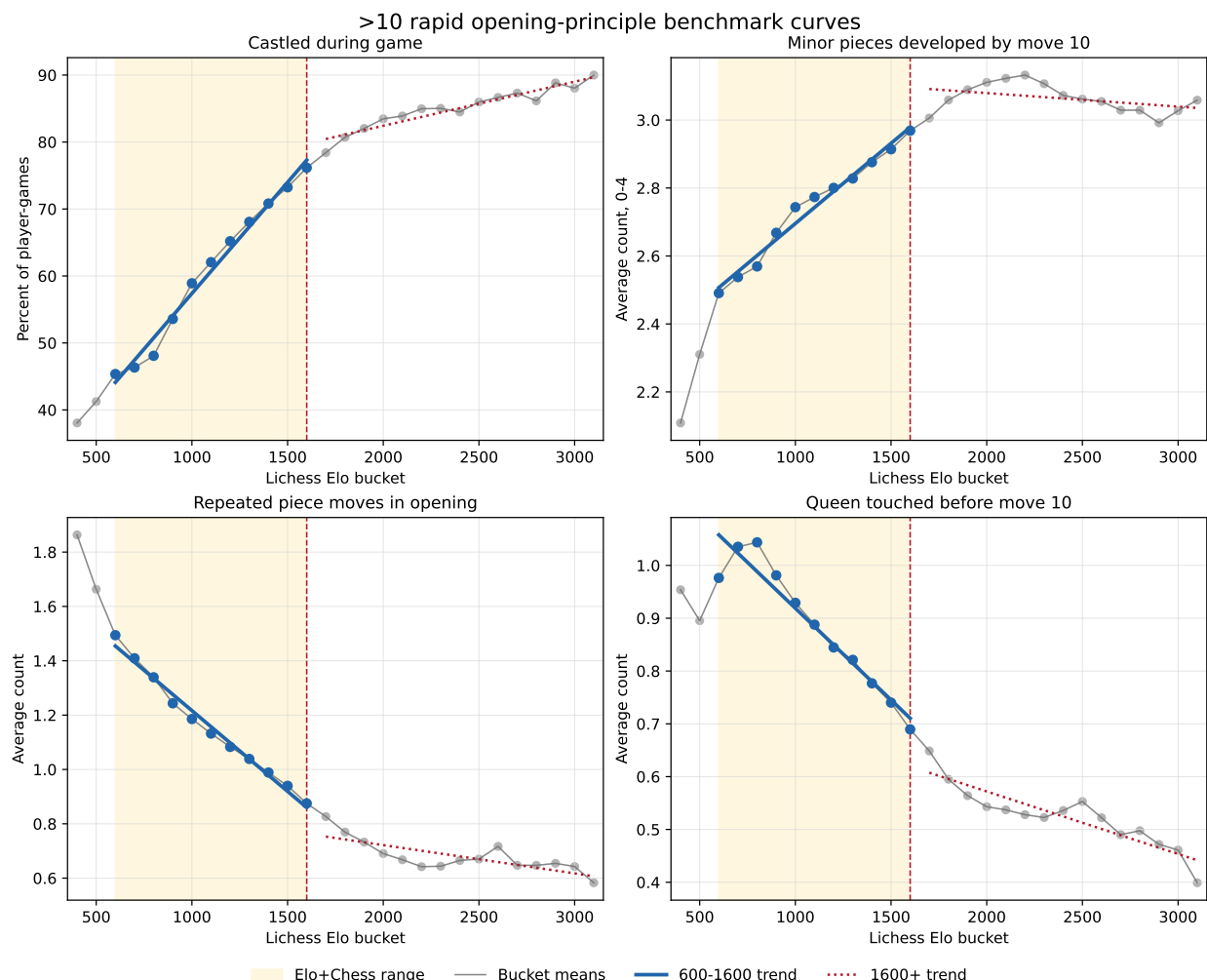


Figure 7: Válogatott nyitási elvű benchmark görbék a hosszabb gyors játékokhoz.

a játékosok közül Elo bucket, majd a legutóbbi kvalifikációs játékokat használja minden egyes kiválasztott játékoshoz. A a jelenlegi célkészlet 11 gyűjtőhelyet használ 600 és 1600 között, 200 játékos / vödör, és játékosonként legfeljebb 50 legutóbbi kvalifikációs játék minden produkcióhoz játék típusa. Ez játéktípusonként 2200 játékos mintát ad, és kb 110 000 játékos-játéksor játéktípusonként a gyűjtőcsoporton belüli százalékosnál disztribúciók.

A 9 ábra szemlélteti a kapott célpontot eloszlás mintavételezett lejátszó konvertált fork rate számára 3+0 blitzben. Mindegyik doboz az gyűjtőkörön belüli eloszlás a mintavételezett játékosok között, nem pedig a gyűjtőkör átlaga egyéni játékok. A jelentés felhasználói felületén a rendszer összehasonlítja a felhasználó értékét a releváns értékkel doboz a felhasználó által feltérképezett Elo vödör számára. Ez az alapja az olyan kijelentéseknek, mint pl hogy egy játékos átváltott vil-larátá magas vagy alacsony-e a társaihoz képest ugyanazon a szinten, még akkor is, ha a teljes lakossági eseményarány az értékeléssel csökken.

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: Hozzávetőleges vödör-görbe meredeksége a négy példa nyitási elvéhez mérőszámok a termelési játéktípusok között.

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: Jelenlegi peer-előzmények százalékos célkészletei. Ezek azok játékos-előzményminták, amelyeket a gyűjtőkörön belüli eloszlások becslésére használnak százalékos stílusú mérőszámok.

12 Replikációs költség

A módszer elvileg reprodukálható, de nem könnyű. Újjáépítés a benchmark görbék megkövetelik:

- hozzáférés a nagy Lichess havi játéktörténet-fájlokhoz;
- rétegzett szűrés játéktípus, értékelési csoport és alkalmasság szerint szabályok;
- legális lépéselemző és bővíthető tábla-állapot-értékelő több tízmillió mozgási állapot;
- jellemző logika a nyitáshoz, anyag, gyalog szerkezet, királybiztonság, bástya tevékenység, taktika, időgazdálkodás és játékeredmények;
- összesítés a mozgás állapotától a játékos-játék jellemzőkig;
- csoportszintű összefoglalók minden jelentésmutatóhoz;
- minden jelölt benchmark görbe manuális áttekintése használat előtt egy edzői jelentést.

3+0 blitznél ez közel félmillió játékot jelent, majdnem egymilliót játékos-játék sorok, és több mint 28 millió mozgás-állapot sor. A négyen át benchmark játéktípusok esetén a jelenlegi benchmark műtermékek több mint 2,55-öt tartalmaznak millió játék, több mint 5,10 millió játékos-játék

Two-sided converted-fork rates by Elo bucket

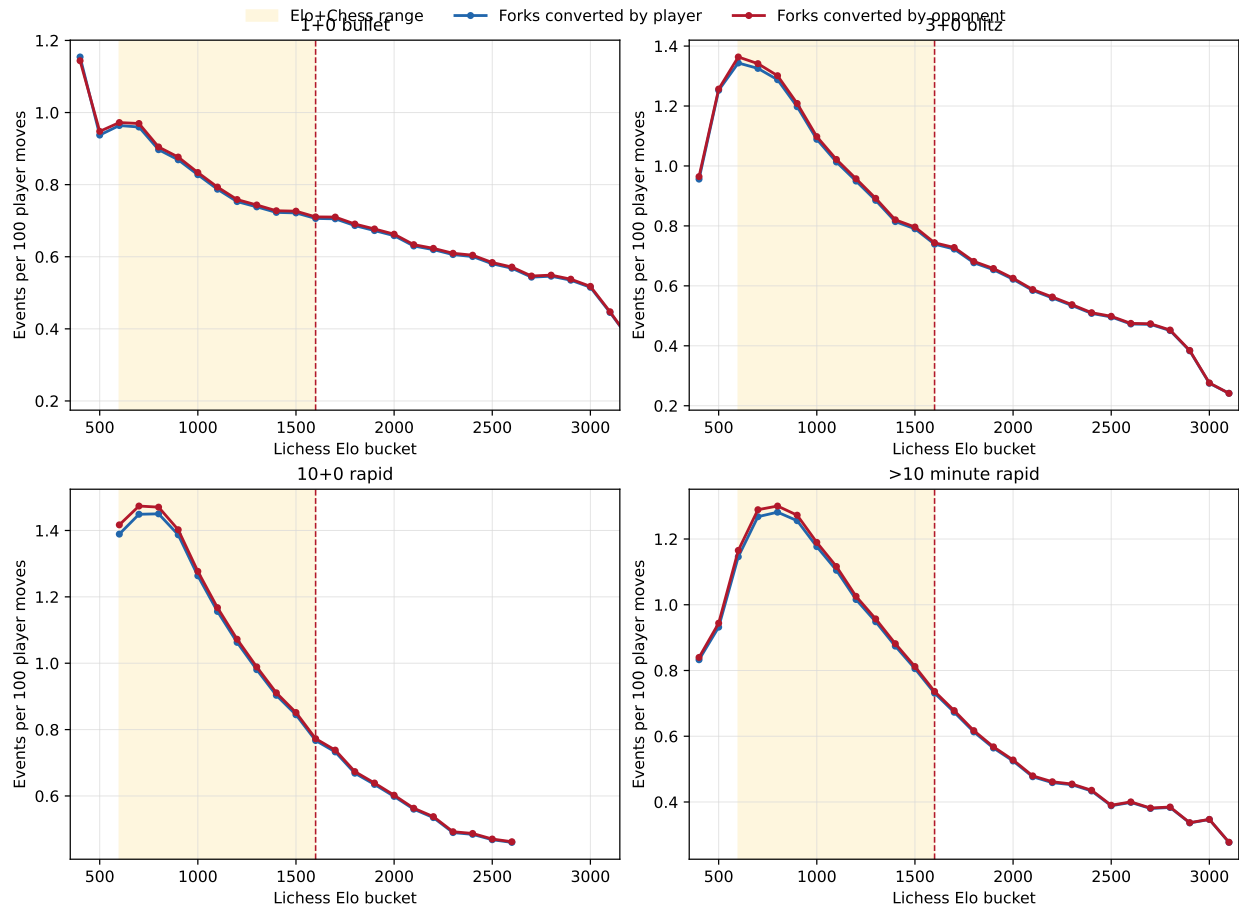


Figure 8: Kétoldalas átalakított villás árfolyamok Elo kanállal. Konvertált villa események esnek magasabb értékeléseknél, mert mindkét fél jobban védekezik és kevesebb tiszta villa konverziók állnak rendelkezésre.

sor és több mint 147 millió mozgás-állapot sor. A skála központi szerepet játszik a benchmark: lehetővé teszi, hogy a jelentés megmutassa a felhasználóknak, mi történik valójában a játszott játékokban a szintjükhöz közel álló játékosok, ahelyett, hogy csak az általános sakkszlogenekre hagyatkoznának.

13 Korlátozások

A benchmark görbék empirikus leírások, nem oksági bizonyítékok. Egy játékos nem szerez értékelést pusztán egy vödör átlagának megfeleltetésével. A mérőszámok is Az összetett pozíciók tökéletlen összefoglalásai: egy korai királynő lépés kiváló lehet az egyik helyzetben, a másikban pedig vakmerő. A görbék célja az azonosítsa a visszatérő szokásokat, és hasonlítsa össze őket a társak viselkedésével, ne pedig azzal helyettesíti a számítást vagy a coaching ítéletet.

A görbék egyben Lichess léptékű görbék is. Amikor a felhasználók hozzák a Chess.com játékok előzmények, a Elo+Chess leképezi a vonatkozó minősítési sávot a Lichess benchmarkra mérete szabható a különálló, többplatformos leképezési modell segítségével. Ez a leképezési réteg

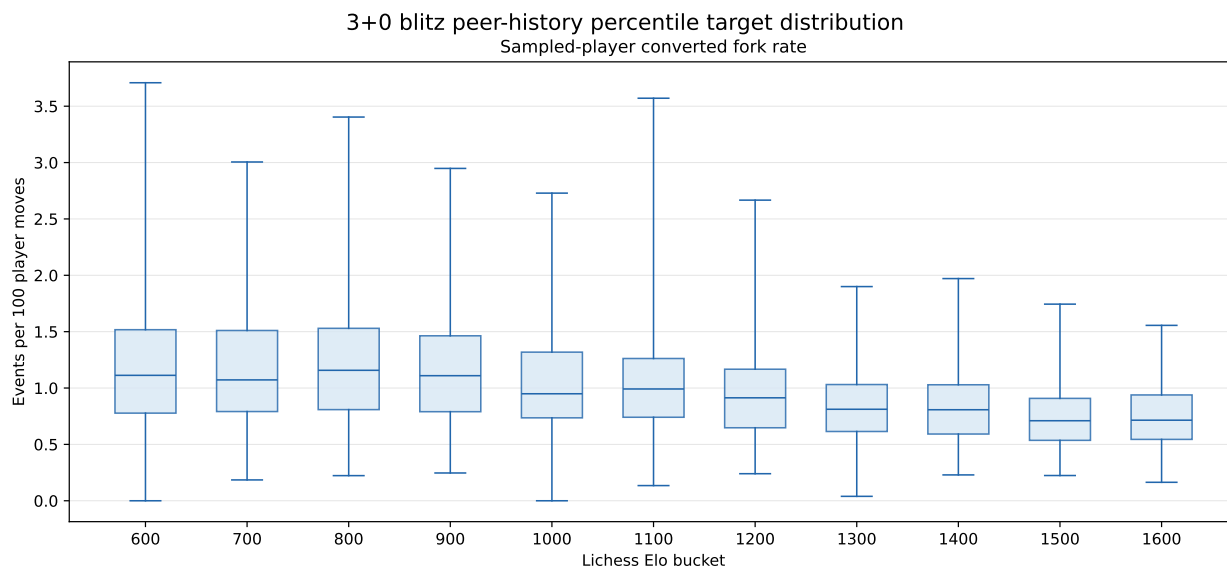


Figure 9: Példa 3+0 blitz peer-history céleloszlás mintavételezett lejátsszóhoz átváltott villa ár-folyam. A százalékos értékeket az Elo csoporton belül számítják ki, így a jelentés összehasonlítja egy játékos olyan társaival, akik nagyjából hasonló taktikai lehetőség előtt állnak árfolyamok.

az a kapcsolódó Elo+Chess kutatási jegyzetben a Lichess-től Chess.com-ig minősítés konverzió: [Estimating Cross-Platform Chess Rating Mappings with Modális regresszió](#). Ez a leképezés nem tekinthető megőrző percentilisnek rangot. Megőrzi a becsült értékelési egyenértékűséget a két minősítés között rendszerek; A százalékos rang platform-specifikus marad, mivel az aktív játékos az eloszlások eltérőek.

14 Gyártási jelentés grafika

A termelési jelentés a benchmark műtermékeket metrika szintű panelekkel alakítja amely a teljes build folyamat ismerete nélkül olvasható. Ábrák [10](#) and [11](#) három példát mutat be az élő Elo+Chess jelentéstervből.

A nyitóelvű panelek a [10](#) ábrán használja a szokásos benchmark-görbe prezentációt. Minden kártya a metrikus család és metrikánév, majd ad egy rövid értelmezési szabályt. A nagy fokozat a jobb felső sarokban szándékosan el van választva a nyeréstől metrikus érték: választ ad a felhasználó által feltett kérdésre, hogy „hogyan értékeli ezt a szokást a benchmarkhoz képest?”, míg az alábbi sorok megőrzik a mért értékeket. A négy összefoglaló sorban az élettartam, az utolsó 10, a nyerések és a veszteségek láthatók. Ez teszi a paneldiagnosztikát, nem pedig pusztán dekoratív. Ha az élettartam és az utolsó 10 megegyezik, a szokás valószínűleg stabil. Ha a győzelmek és a veszteségek eltérnek, a mérőszám lehet eredményfüggő vagy játékállapothoz kötött, nem pedig egyszerű szokás.

A diagram terület visszafogott sötét háttérrel, arany trendvonalat és színeset használ jelölők a négy szelethez. A trend annotációja kifejezetten kimondja, hogy a benchmark kapcsolat emelkedik vagy süllyed Elóval. Ez azért fontos, mert egyes mutatók magasabb-is-jobb, mint például a kisebb darabok fejlesztése, míg mások az alacsonyabb-is-jobb, mint például a korai királynő mozgás. A kártya tehát nem kéri a felhasználót, hogy egyedül a lejtőből következtessen irányra. A döntő A „Mit

jelent ez” és a „Coaching tipp” sorok lefordítják a benchmark eredményét konkrét sakk-oktatásba az empirikus kontextus megőrzése mellett.

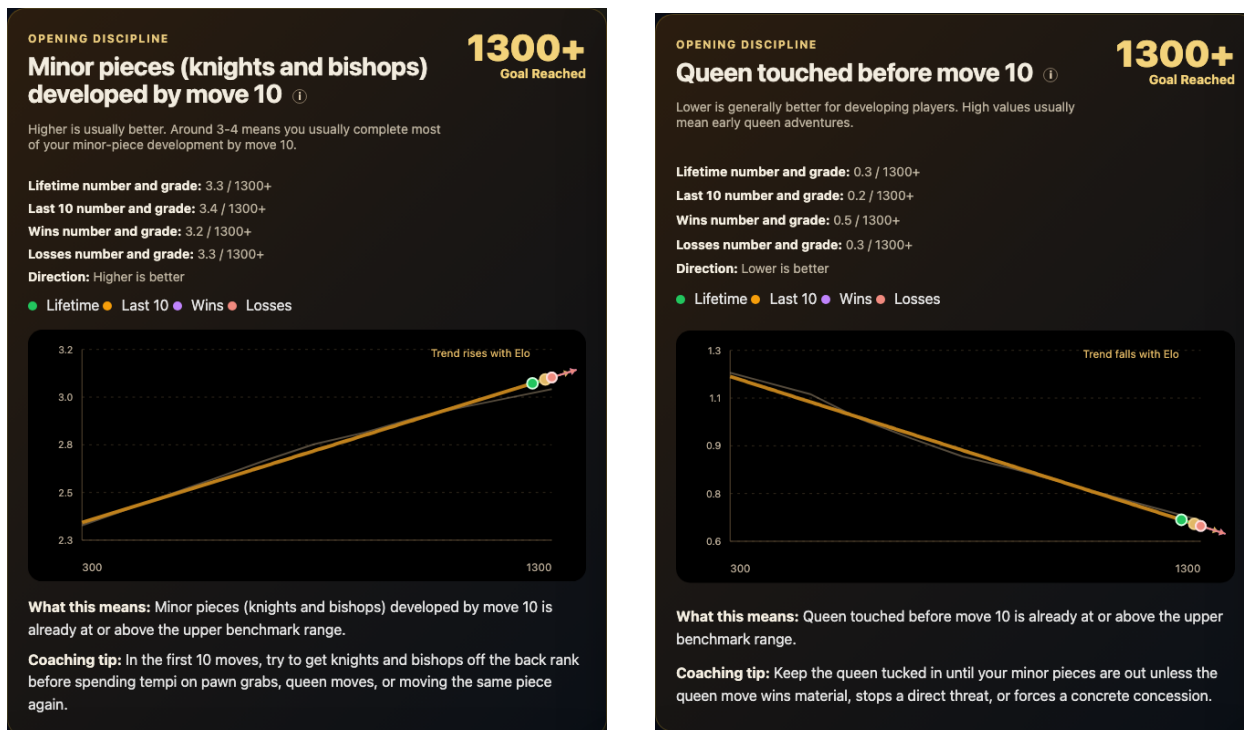


Figure 10: Gyártási nyitóelvű panelek. Ugyanez a vizuális nyelvtan kezeli a magasabb a jobb metrika, a kisebb darabok fejlesztése és a kisebb a jobb mutató, korai királynő mozgalom. Mindkét panelen láthatók az élettartam, az utolsó 10, a nyerések és a veszteségek így a jelentés elkülönítheti a stabil szokásokat a közelmúlttól vagy az eredménytől függőtől minták.

A 11 ábra a percentilis stílust mutatja olyan metrikákhoz használt prezentáció, amelyek nyers arányát nem reprezentálja jól a egyetlen globális trend. A nyerő villa egy taktikai példa: a népszerűség aránya A konvertált villák száma a játékos képességeivel és az ellenfél hibaarányával egyaránt változik. A A kártya ezért összehasonlítja a felhasználó értékét a játékostársak történeteivel feltérképezett Elo vödör. A boxplot megjeleníti a peer eloszlást, beleértve a interkvartilis tartomány, medián és bajusz. A felhasználó markere a ugyanaz a skála, és mind a nyers értékkel, mind pedig a percentilis ranggal címkézve. Szomszédos Az összehasonlító gyűjtők némitott formában láthatók maradnak, így a felhasználó láthatja, hogy a peer vödör helyi, nem univerzális skála.

Ez a kialakítás auditálhatóvá teszi a percentilis mérőszámokat. A felhasználó látja a nyers értéket, a százalékos, az összehasonlításához használt peer vödör és a lefordított Chess.com-Lichess vödör leképezés. Ugyanez élettartam, utolsó 10, nyerések és A veszteségek sorai megmaradnak, így a taktikai százalékpontok továbbra is ellenőrizhetők az aktuális hatások és az eredményfüggőség miatt.

15 Következtetés

A Elo+Chess benchmark jelentések egy nagy mozgási állapot számításból épülnek fel rétegzett Lichess minták. Minden mérőszám konkrét táblaállapot halmazával kezdődik és mozgásállapot

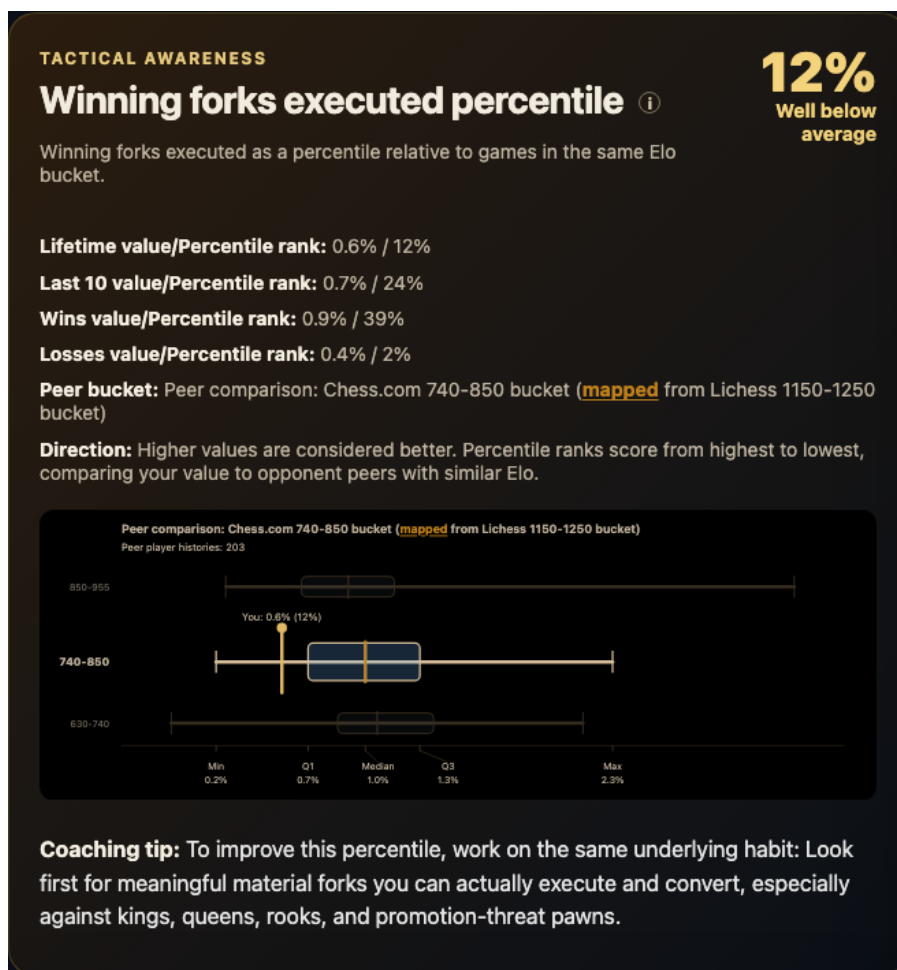


Figure 11: Gyártási percentilis panel a nyertes villák végrehajtásához. A boxplot a jelentés által használt azonos csoportos társeloszlást mutatja, míg az aranyat marker a felhasználó nyers értékét és százalékos rangját mutatja.

változók, játékos-játék jellemzővé válik, majd átlagolják Elo vödör segítségével empirikus benchmark görbét alkotva. A nyitás elve példák mutatják meg, miért hasznos ez a megközelítés: az egyszerű sakkszokások erősek, látható kapcsolatok értékelésével a céltartományban, míg azok a kapcsolatok gyakran ellaposodnak vagy feltételesebbé válnak a felső határ felett. Ez a központi oka annak, hogy a Elo+Chess coaching állításait a minősítésre összpontosítja ahol az adatok a legvilágosabban alátámasztják őket.