

Elo+Chess에 대한 계층화된 Lichess 벤치마크 곡선 구축

Elo+Chess 연구 노트

2026년 5월 31일

Abstract

Elo+Chess 벤치마크 보고서는 대규모 계층화된 샘플로 구성됩니다. Lichess.org 게임(이 노트에 사용된 벤치마크 샘플 포함) 1+0 블릿 게임 483,974개, 3+0 블리츠 게임 479,974개, 10+0 래피드 게임 925,636개, 660,540개의 더 길고 빠른 게임. 각 주요 게임 유형에 대해 파이프라인이 확장됩니다. 모든 게임을 이동 전 및 이동 후 보드 상태로 전환하고 수백 가지를 계산합니다. 중간 및 플레이어 게임 변수는 해당 변수를 코칭 측정항목을 사용한 다음 Lichess Elo 버킷을 기준으로 해당 측정항목의 평균을 구합니다. 는 샘플은 원시 플레이어에 비례하지 않고 의도적으로 계층화됩니다. 폴: 목적은 각 등급 수준에서 안정적인 행동 곡선을 추정하는 것입니다. 자연적인 모습을 재현하지 않기 위해 드물게 사용되는 버킷도 포함합니다. 평점 분포. 그 결과는 경험적 벤치마크 곡선군입니다. 각 측정항목을 통해 서로 다른 Elo 레벨의 플레이어가 주어진 수치를 얼마나 자주 표시하는지 물어볼 수 있습니다. 실제 게임의 습관이나 원칙. 이 노트에서는 방법, 규모에 대해 설명합니다. 계산, 그리고 여러 캐슬링, 마이너 피스 개발, 반복되는 오프닝 피스 이동 및 초기 퀸 이동. 강조점은 방법론: Elo+Chess는 체스 원칙이 보편적이라고 가정하지 않습니다. 모든 기술 수준에 걸쳐 선형입니다. 대신 벤치마크 곡선은 다음을 통해 검사됩니다. 등급 범위 및 제품은 초급부터 초기 고급 영역에 중점을 둡니다. 관계가 가장 강력하고 가장 해석하기 쉬운 곳입니다.

1 목적

벤치마크 시스템의 목표는 광범위한 체스 조언을 다음으로 변환하는 것입니다. 측정 가능한 수량. '당신의 작은 작품을 개발하세요'와 같은 진술은 유용하지만, 하나의 작은 조각이 몇 개인지 말할 수 있으면 더 실행 가능해집니다. 플레이어는 일반적으로 10번째 행으로 발전하는데, 이는 동일한 동료와 비교하여 어떻습니까? 등급 버킷, 등급 스펙트럼 전반에 걸쳐 습관이 어떻게 변하는 지.

따라서 Elo+Chess는 실제 게임에서 메트릭 라이브러리를 구축합니다. 각 지표는 게임 유형 및 Elo 버킷별로 요약된 모든 플레이어-게임 행에 대해 계산됩니다. 보고서의 벤치마크 곡선으로 사용됩니다. 그런 다음 보고서는 사용자를 다음 위치에 배치할 수 있습니다. 동일한 곡선을 사용하여 사용자가 행동의 위, 아래 또는 근처에 있는지 표시합니다. 비슷한 플레이어들 사이에서 관찰되었습니다.

2 소스 데이터 및 계층화

벤치마크 소스는 공개 Lichess.org 게임 데이터베이스입니다. Elo+Chess는 다음을 사용합니다. 소규모보다는 게임 유형 및 등급 버킷별로 큰 계층화된 샘플 편의 샘플. 현재 프로덕션 런타임 아티팩트는 네 가지 주요 내용을 다루고 있습니다. 게임 카테고리:

- 1+0 총알,
- 3+0 블리츠,

- 10+0 급속,
- 10분 이상의 빠른 게임.

계층화 목표는 등급 수준 전반에 걸친 폭입니다. 비례 무작위 샘플은 전체 플레이어 풀을 설명하는 데 효율적이지만 벤치마크 곡선을 구축하는 데 비효율적입니다. 가장 일반적인 평가 버킷 샘플을 지배하는 반면 지지력이 낮은 테일 버킷은 너무 적습니다. 안정적인 측정 평균을 위한 게임. 따라서 Elo+Chess는 평균 게임별로 샘플링합니다. 엘로 버킷. 현재의 클린 샘플 구성에서는 원시 게임이 먼저입니다. 알려진 플레이어, 알려진 등급, 유효한 결과가 있는 유효한 등급의 게임으로 필터링되었습니다. 그리고 알려진 시간 제어. 그런 다음 게임은 충분한 게임을 가진 플레이어로 필터링됩니다. 일관성 있는 기술 수준을 나타낼 만큼 속도와 안정성이 충분한 등급입니다. 각 플레이어에 대해 Elo+Chess는 플레이어의 속도 내 등급 범위를 계산합니다. $(\max(\text{Elo}) - \min(\text{Elo}))$ 를 소스 풀 전체에 걸쳐 유지하고 두 플레이어 모두 최소한 구성된 최소 게임 수를 보유하고 있는 경우에만 게임을 진행합니다. 속도 내 드리프트의 Elo 포인트는 300점을 넘지 않습니다. 드리프트 필터는 의미가 없습니다. 정상적인 개선이나 변동을 제거하기 위해; 플레이어가 게임은 임시 계정, 복귀 플레이어 또는 여러 버킷을 통해 빠르게 이동한 계정 샘플링 기간. 마지막으로 적격 게임은 각 평균 Elo 내에서 순위가 매겨집니다. 게임 ID와 시드의 결정적 해시와 제한된 수의 버킷 게임은 각 버킷에 보관됩니다.

이 디자인은 벤치마크 곡선을 훨씬 더 유용하게 만듭니다. 각 Elo 버킷 각 지표의 평균값을 추정하기에 충분한 관찰 결과를 제공합니다. 희귀한 높거나 낮은 등급의 버킷은 자연스러운 모양에 의해 익사되지 않습니다. 활성 플레이어 배포의. Elo+Chess에서 사용되는 벤치마크 곡선은 다음과 같습니다. 그런 다음 제품과 가장 관련성이 높은 범위로 제한됩니다. 대략 초보자입니다. 초기 고급 플레이를 통해. Lichess 규모 보고서에서 해당 기본 범위는 600 1600입니다. 사용자가 Chess.com 척도로 보고서를 볼 때, 해당 평가 버킷은 동일한 기본 Lichess에 매핑됩니다. 별도의 크로스 플랫폼 등급 매핑 방법을 사용하는 벤치마크 버킷 동반 논문에 설명되어 있습니다: [Estimating Cross-Platform Chess Rating Mappings with 모달 회귀](#).

3 활성 플레이어 등급 분포

종료 시점의 Lichess.org 플레이어 Elo 등급 분포를 살펴보면 2026년 3월 기준 주요 게임 유형별로 다음과 같은 개수와 백분위수가 표시됩니다. 이 분포 보기의 경우 각 플레이어는 정확한 게임 유형별로 한 번씩 계산됩니다. 해당 원시에서 해당 플레이어의 최근 관찰된 등급을 사용합니다. 해당 게임에서 최소 5번의 게임을 요구한 후 시간 제어 행 유형.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: 최근 관찰된 현역 플레이어 Lichess 등급 분포를 정확하게 해당 게임 유형에서 최소 5개의 게임이 필요한 후 게임 유형을 선택하세요.

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 이 필터가 있는 분포와 없는 분포를 표시합니다.

이 분포를 계층화된 벤치마크 샘플과 혼동해서는 안 됩니다. 분포는 원시에서 관찰된 활성 Lichess 플레이어 풀을 설명합니다. 스캔. 그런 다음 벤치마크 샘플은 의도적으로 계층화되어 Elo 버킷이 안정적인 미터법 곡선을 충분히 지원합니다.

특정 게임 유형에 대한 Lichess.org Elo 등급을 다음으로 매핑하려면 동등한 Chess.com 등급, Elo+Chess는 게임 유형별 모달을 사용합니다. 컴패니언에서 추정된 회귀 방정식 [Elo+Chess rating-mapping paper](#):

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

이러한 방정식은 백분위수 순위가 아닌 등급 값을 매핑합니다. 백분위 순위는 특정 플레이어 풀 내부에 정의되어 있으며 Chess.com 및 Lichess 플레이어 수영장은 모양이 다양해요. 테이블 2에서 10분 빠른 게임 유형, 첫 번째 열은 가능한 작은 샘플을 제공합니다. Chess.com Elo 등급 값. 두 번째 열은 해당 항목을 제공합니다. 2026년 5월 31일 현재 Chess.com에 표시된 Chess.com 백분위수 순위. A 71.7

10분 빠른 속도로 Lichess.org 순위는 Lichess Elo 등급에 해당합니다. 1,644²이며 해당 값은 세 번째 열. Chess.com-Lichess 매핑 방정식 1,644를 적용하면 Lichess는 Chess.com의 약 1,271에 해당합니다. 추정을 가정하면 매핑 방정식이 합리적이면 Chess.com 등급 사이의 +522 간격이 열 1과 열 4의 매핑된 등급은 큰 숫자로 설명될 수 있습니다. Chess.com를 팽창시키는 Chess.com 플레이어 풀의 캐주얼 플레이어 백분위 순위. 즉, Lichess.org에서 72번째 백분위수에 속한다는 것은 Chess.com에서 72번째 백분위수에 속하는 것보다 더 큰 성과입니다. 백분위수 순위는 한 시스템에서 다른 시스템으로 명확하게 매핑되지 않습니다. 실제 엘로 값은 동일한 플레이어 점수에서 추정된 방정식을 사용하여 매핑될 수 있습니다. 두 시스템.

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: Chess.com 10분 빠른 백분위수 비교 예시 활성 Lichess 10+0 rapid에서 포인트와 동일한 백분위수 포인트 최소 5개의 10+0 게임이 필요한 후 배포됩니다. “매핑된 Chess.com Equiv.”는 현재 Elo+Chess를 적용합니다. 10+0 등급 규모를 동일한 백분위수 Lichess 등급으로 빠르게 변환합니다. 는 눈금 구별을 표시하기 위해 가장 높은 꼬리 행이 포함되었지만 적합 등급 라인은 주요 코칭 범위에서 가장 신뢰할 수 있습니다.

Table 2에 공백이 있는 이유는 매핑이 동등한 연주 강도 등급을 추정하는 반면, 백분위수는 하나의 플랫폼 자체 인구 내에서 위치. 71.7번째 백분위수 플레이어 따라서 Chess.com는 Lichess에서 자동으로 71.7번째 백분위수 플레이어가 아닙니다. 그 반대도 마찬가지입니다. 관찰된 수치는 Chess.com의 공개 수치와 일치합니다. 낮은 등급 또는 캐주얼 플레이어 질량이 훨씬 더 큰 빠른 풀 활성 정밀 제어 Lichess 풀은 숙련된 반복자에게 더 집중되어 있습니다. 플레이어. 중요한 점은 하나의 백분위수 척도가 ‘옳다’는 것이 아니라 다른 하나는 ‘잘못’입니다. 백분위 순위는 플랫폼 간에 이식 가능하지 않습니다. 기본 플레이어 분포가 동일한 모양을 갖지 않는 한.

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.

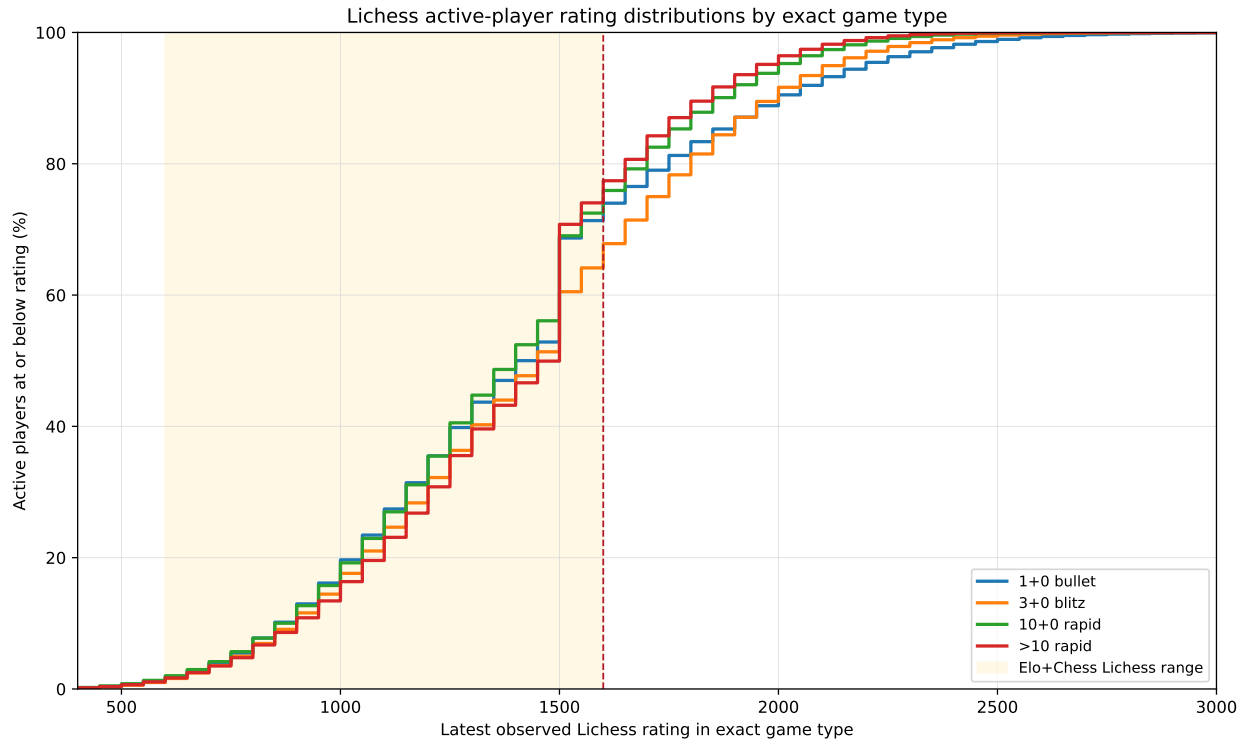


Figure 1: 정확한 게임 유형별 누적 활성 플레이어 등급 분포입니다. 는 노란색 띠는 기본 Elo+Chess 에서 사용되는 600 1600 Lichess 범위를 표시합니다. 코칭 벤치마크.

Note on the 1500 spike. The visible jump near 1500은 스무딩에 의해 생성된 것이 아니라 데이터에 존재합니다. 최근에는 관찰된 평점 분포, 정확한 1500개의 평점은 1+0의 13.1

Bullet 플레이어, 3+0 Blitz 플레이어의 5.6

17.6

등급 시스템의 시작 또는 임시 앵커

정확히 1500 스파이크는 주로 활동이 적은 계정의 인공물입니다. 받는 사람 확인해 보세요. 요구한 후 동일한 최신 등급 분포를 다시 계산했습니다. 각 플레이어는 정확한 게임 유형에서 최소 5개의 게임을 플레이해야 합니다. 스파이크가 거의 이 필터에서는 사라집니다. 정확한 1500 등급은 5.6-17.6

5개 이상의 게임에서 전체 플레이어 분포의 플레이어 수는 0.2 1.2

Table 1에 보고된 분포.

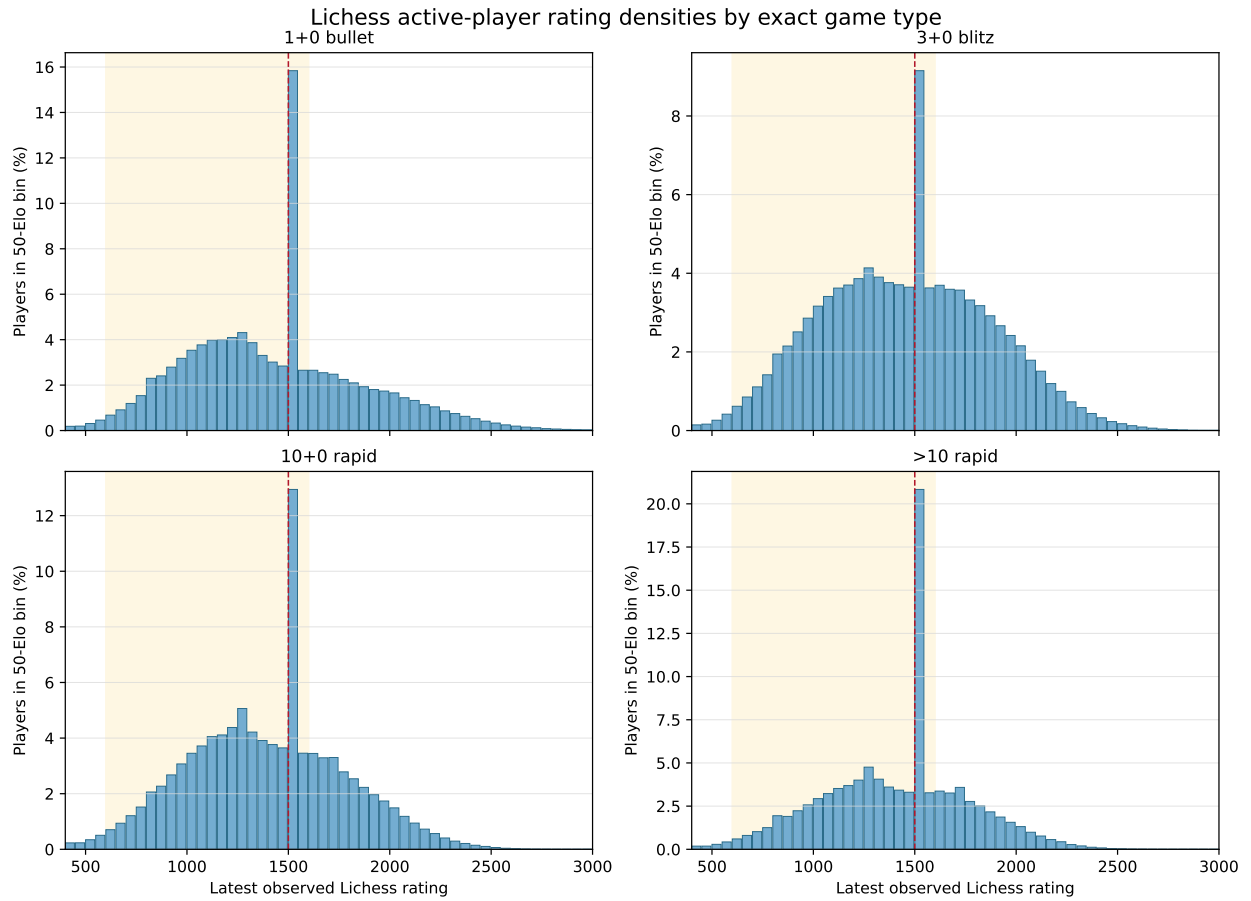


Figure 2: 동일한 활성 플레이어 등급 분포의 밀도 보기. 는 점선은 1500 Lichess를 표시하며, 여기서 원시 최신 등급 분포는 큰 정확한 등급 질량 포인트. 노란색 밴드는 600–1600 Lichess를 표시합니다. 기본 Elo+Chess 측정 보고서에서 사용되는 범위입니다. 스파이크가 가장 눈에 띄는 것은 활동이 적은 많은 계정이 정확하게 남아 있기 때문에 전체 플레이어 밀도 1500; Figure 4는 밀도 플롯을 반복합니다. 정확한 게임 유형에서 최소 5개의 게임을 요구한 후 이 유물.

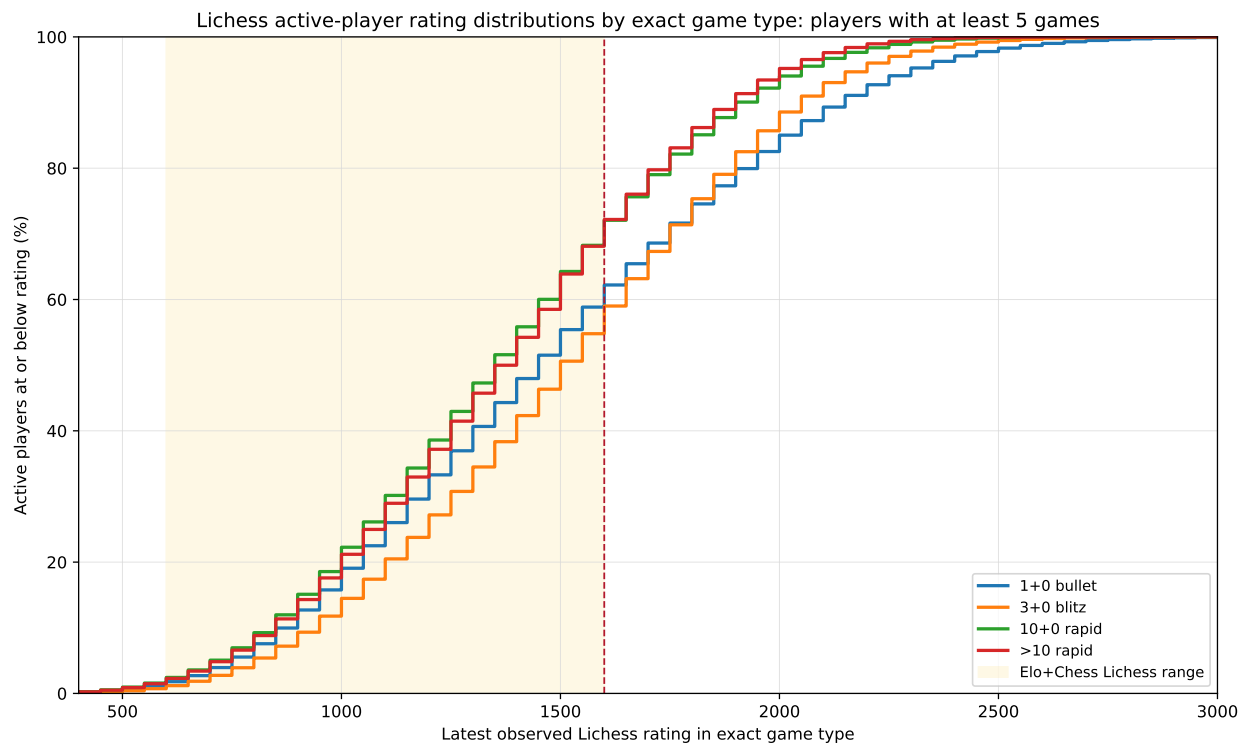


Figure 3: 요구 후 누적 활성 플레이어 등급 분포 정확한 게임 유형의 게임이 최소 5개 이상이어야 합니다. 1500점프가 대폭 감소됩니다.

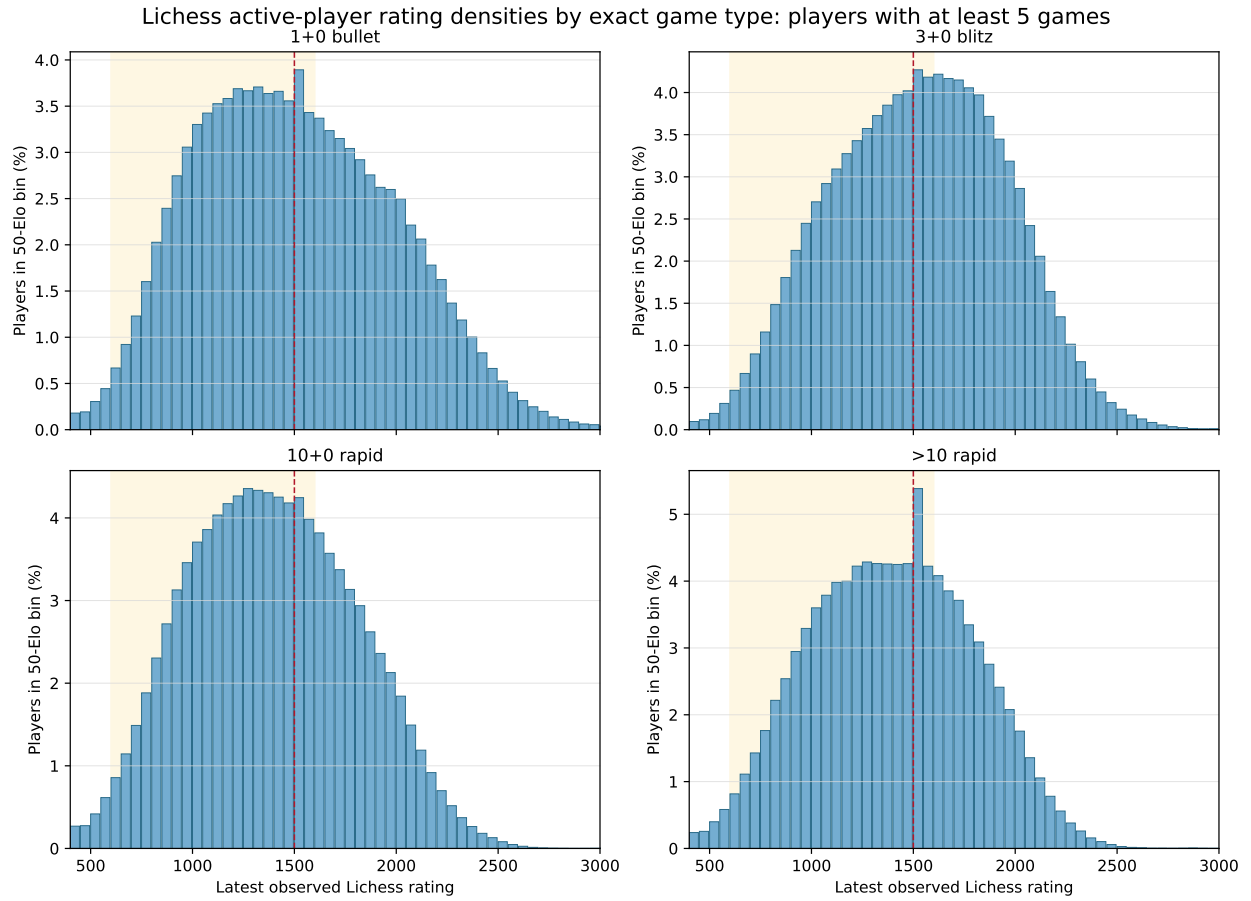


Figure 4: 해당 게임에서 최소 5개의 게임을 요구한 후 밀도 보기 유형. 이 견해는 올 플레이어에서 대규모 1500 매스 포인트가 있음을 확인합니다. 배포는 주로 활동이 적은 계정에 의해 이루어집니다. 노란색 밴드 표시 기본 Elo+Chess 지표 보고서에서 사용되는 600 1600 Lichess 범위.

4 현재 벤치마크 아티팩트의 규모

계산량이 의도적으로 큼니다. 프로필 굵은 자국이 아니고 작은 다운로드한 PGN 세트 또는 직접 선별한 유익한 예제 컬렉션입니다. 초기 Lichess 패스에는 2025년 1월부터 전체 월별 파일이 포함됩니다. 계층화되지 않은 스캔 테이블에는 1,382,178,087개의 게임이 포함되어 있습니다. 플레이어-게임 등급 행은 2,764,356,174개입니다. 해당 행 전체에서 6,623,438개의 고유한 플레이어 사용자 이름이 한 번 이상 나타납니다.

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: 벤치마크 샘플링 전 계층화되지 않은 Lichess 스캔 규모. 스캔 2025년 1월부터 3월까지의 전체 월별 게임 기록 파일을 다룹니다. 2026.

프로덕션 벤치마크 아티팩트는 원시 스캔보다 훨씬 작습니다. 미터법 구성을 위해 선택됩니다. 여전히 규모가 크다: 벤치마크 단계는 수백 개의 이동 상태 확장 및 집계 파이프라인입니다. 게임 유형당 수천 개의 게임.

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: 정렬 후 계층화된 벤치마크 아티팩트의 규모 일반적인 277열 런타임 기능 스키마. 에이 '플레이어-게임 행'은 한 플레이어의 관점에서 본 하나의 게임을 의미하므로 대부분의 게임은 두 개의 플레이어-게임 행을 생성합니다.

3+0 블리츠의 경우에만 현재 프로덕션 벤치마크에서는 479,974개의 게임을 사용합니다. 129,398명의 고유 플레이어와 28,138,904개의 개별 이동 상태 행이 포함됩니다. 는 정렬된 블리츠 플레이어-게임 기능 테이블에는 277개의 파생 변수가 포함되어 있습니다. 는 이동 상태 단계에서는 신속한 기능 구축 아티팩트가 훨씬 더 넓어집니다. 기능 구성 중에 사용되는 빠른 이동 상태 테이블에는 168개가 포함됩니다. 277열 플레이어-게임 기능으로 집계되기 전의 이동 상태 열 테이블. 네 가지 게임 유형 모두에 대한 프런트엔드 런타임 벤치마크 아티팩트는 다음과 같습니다. 그런 다음 동일한 277열 플레이어-게임 기능 스키마에 맞춰 정렬됩니다. 전술-분류 브레이크아웃 변수가 확장되었습니다. 이들 특징 테이블로부터, 보고서는 가장 해석하기 쉬운 코칭 지표를 선택하고 표시합니다.

이러한 개수는 결과를 해석하는 데 중요합니다. 벤치마크 곡선 이는 수동으로 검토한 수백 개의 게임에 대한 추정치가 아닙니다. 대신에 그들은 수천만 개의 평가를 받은 이사회와 4개 이상의 주에서 왔습니다. 현재 4개의 프로덕션 게임 카테고리에 걸쳐 수백만 개의 플레이어-게임 행이 있습니다.

5 파이프라인 개요

벤치마크 구축 파이프라인에는 6가지 주요 단계가 있습니다.

1. 대상 게임 유형 및 등급 계층에서 등급 Lichess 게임을 선택합니다.
2. 각 게임 동작을 동작별로 분석하고 동작 전과 동작 후를 저장합니다. 보드 상태.
3. 보드에서 중간 이동 상태 변수를 계산합니다. 플레이어 관점, 상대 관점.
4. 이동 상태 변수를 플레이어 게임 기능 변수로 집계합니다.
5. 플레이어-게임 변수를 명확한 보고서 측정항목으로 변환합니다. 해석 방향: 일반적으로 높을수록 좋고 낮을수록 좋습니다. 더 좋거나 설명적입니다.
6. Elo 버킷별로 각 지표를 평균화하고 결과를 수동으로 검사합니다. 관계.

많은 보고서 측정항목이 직접적으로 측정되지 않기 때문에 이 구조는 중요합니다. PGN 헤더에 표시됩니다. 위치를 재구성하고 결정해야 합니다. 어떤 조각이 어느 시작 사각형에서 이동했는지, 이동 전과 비교 이동 후 재료 및 구조를 수집한 후 그 결과를 집계합니다. 레벨을 플레이어-게임 레벨로 이동합니다.

6 중간변수

중간 변수는 보드에서 변경된 사항을 설명하도록 설계되었습니다. 신고자 관점에서 이것이 왜 중요한지 알아보세요. 예는 다음과 같습니다:

- 이동 번호, 플레이어 이동 번호, 이동할 면, 무버 색상 및 플레이어 색상;
- 정사각형, 정사각형, 이동 조각 유형, 캡처 조각 유형 및 이동이 캡처, 확인, 성, 승진 또는 짝인지 여부;
- 이동 전 및 이동 후 재료 균형;
- 작은 조각이 주어진 이동 수만큼 시작 사각형을 벗어났는지 여부;
- 연속적인 플레이어가 같은 조각을 이동했는지 여부 열기;
- 퀸이 10번째 행 이전에 움직였는지 여부;
- 캐슬링 쪽, 캐슬링 이동 수, 룯이 연결되었는지 여부 캐슬링 후;
- 느슨한 조각, 매달린 조각, 이중 폰, 격리된 폰, 통과된 폰 및 뒤로 폰;
- 루크 파일 상태, 열린 파일, 반쯤 열린 파일 및 킹 안전 특징;
- 전술적 기회, 실행된 전술, 물질적 압박 전술, 후속 이익 보장 및 실현 이익.

정확한 중간 세트는 빌드 단계에 따라 다릅니다. 런타임 보고서 데이터베이스는 보고서를 신속하게 제공하는 데 필요한 열만 포함하고 전체 기능 구축은 아티팩트에는 더 넓은 이동 상태 테이블과 구성에 사용되는 기능 패치가 포함됩니다. 마지막 플레이어-게임 기능 행.

7 개방 원리 측정법의 예

개방 원리는 사용자가 쉽기 때문에 유용한 예입니다. 이동 상태 계산을 이해하고 쉽게 연결할 수 있습니다.

7.1 Minor Pieces Developed by Move 10

이 측정항목은 플레이어의 4명의 기사와 주교 중 얼마나 많은 사람이 남았는지 계산합니다. 플레이어의 10번째 이동으로 흠 스쿼어가 됩니다. 중간 계산 무버 색상, 이동 조각 유형, 시작 사각형 및 플레이어 이동을 사용합니다. 번호. 백의 경우 흠 스쿼어는 b1, g1, c1 및 f1입니다. 블랙의 경우, b8, g8, c8, f8. 각 플레이어-게임에 대해 Elo+Chess는 고유한 흠을 계산합니다. 플레이어 이동에 따라 이동한 기사와 비숍 사이의 사각형 10.

이는 일반적으로 높을수록 더 나은 측정항목입니다. 모든 미성년자가 그렇다고 주장하는 것은 아닙니다. 조각은 항상 10번째 움직임으로 움직여야 하지만, 발전하는 플레이어들 사이에서는 경험적 관계는 분명합니다. 더 높은 Elo 버킷에 도달한 플레이어가 발전하는 경향이 있습니다. 더 일찍 작은 조각.

7.2 Repeated Piece Moves in the Opening

이 측정항목은 동일한 조각을 다시 이동하는 데 소요된 추가 초기 이동을 계산합니다. 새로운 것을 개발하는 것. 중간 계산에서는 이전 계산을 사용합니다. 대상 사각형, 현재 원점 사각형 및 플레이어 이동 번호입니다. 만약에 이전 플레이어가 움직일 때 같은 조각이 움직였고 그 안에서 다시 움직였습니다. 처음 10번의 플레이어 이동이 완료되면 반복 이동 횟수가 증가합니다.

이는 일반적으로 낮을수록 더 나은 측정항목입니다. 그것은 낭비되는 개방 온도를 측정합니다. 간단한 방법. 일부 반복적인 움직임은 전술적으로 정당하지만 규모에 따라 버킷 곡선은 일반적인 경향을 포착합니다. 즉, 목표에 더 강한 플레이어가 있는 것입니다. 레벨은 동일한 조각을 반복적으로 이동하는 데 더 적은 초기 동작을 소비합니다.

7.3 Queen Touched Before Move 10

이 측정항목은 플레이어 이동 10 이전의 여왕 이동을 계산합니다. 중간 계산에서는 이동 조각 유형과 플레이어 이동 수를 사용합니다. 움직이는 조각이라면 퀸이고 플레이어 이동 횟수가 최대 10일 경우 카운트가 증가합니다.

이는 Elo+Chess 목표 범위의 경우 일반적으로 낮을수록 좋습니다. 초기 여왕의 움직임은 물질적 양보를 얻거나 강제적 양보를 얻을 수 있지만, 많은 개발 중인 플레이어는 여왕을 너무 일찍 데려오고, 여왕에 대한 공격에 대한 템피를 잃고, 지연됩니다. 마이너 피스 개발 및 킹 세이프티.

7.4 Castling and King Safety

캐슬링 지표는 캐슬링 플래그, 캐슬링 측면, 캐슬링 이동 수 및 캐슬링 후 보드 상태. 예를 들면 플레이어가 캐슬링했는지 여부, 플레이어가 10번째 수로 캐슬링했는지 여부, 플레이어가 킹사이드에 캐슬링했는지 여부 퀸사이드, 캐슬링 후 룯이 연결되었는지, 룯 파일이 연결되었는지 여부 캐슬링 후에는 열려 있거나 반쯤 열려 있었습니다.

이러한 측정항목은 모두 동일한 방식으로 해석되지 않습니다. 캐슬링 완성 및 이동 10으로 캐슬링하는 것은 일반적으로 목표 범위에서 높을수록 좋습니다. 퀸사이드 캐슬링이나 열린 파일로의 캐슬링은 상황에 따라 더 달라질 수 있습니다. 이것이 바로 Elo+Chess가 단순히 도덕적 규칙을 하드 코딩하지 않는 이유입니다. 버킷을 만듭니다 어떤 관계가 코칭에 충분히 안정적인지 곡선을 그리고 검사합니다.

8 버킷 곡선 및 수동 검사

플레이어-게임 측정항목이 계산된 후 Elo+Chess는 이를 Lichess Elo별로 그룹화합니다. 버킷을 만들고 버킷 평균을 계산합니다. 버킷 평균은 원시 경험적입니다. 해당 측정항목의 벤치마크 포인트입니다. 그런 다음 보고서에서 추세선을 사용하여 다음을 수행합니다. 사용자 값에 Elo 스타일 등급을 할당합니다.

곡선은 사용하기 전에 수동으로 검사됩니다. 측정항목이 더 유용합니다. 버킷이 등급 범위 전반에 걸쳐 일관된 관계를 표시하는 것을 의미하는 경우 제품으로 제공됩니다. 측정항목은 잡음이 많고, 편평하며, 매우 높을 때 덜 유용합니다. 단조롭지 않거나 사용자에게 설명하기 어려운 특수한 경우에 의해 구동됩니다.

이 수동 검사 단계는 Elo+Chess 갑피 위에서 특히 중요합니다. 컷오프. 더 강한 플레이어는 언제 오프닝 원칙을 깨야 하는지 알고 그렇게 합니다. 효과적으로. 초급부터 고급까지의 범위 이상에서는 관계가 단순한 습관과 평가 사이에서 종종 평탄화되거나 형태가 변합니다. 제품 따라서 경험적 관계가 강한 범위를 강조합니다. 교육적으로도 유용합니다.

9 전대역 개방 곡선의 예

그림 5는 4개의 3+0 블리츠 개방 원리 곡선을 보여줍니다. 회색 점은 더 넓은 Lichess 등급 범위를 나타냅니다. 블루포인트와 블루트렌드 라인은 메인 Elo+Chess 코칭에서 사용되는 600 1600 Lichess 범위를 보여줍니다. 벤치마크. 빨간색 점선은 상위 컷오프를 표시합니다.

동일한 패턴이 전체 범위 10+0 급속 및 장거리 고속에서 볼 수 있습니다. 유물. 중앙 모델링 결정은 변경되지 않았습니다. 경험적으로 사용하십시오. 코칭 주장의 안정적인 범위와 단순한 습관 곡선인 척하는 것을 피하십시오. 영원히 똑같이 유익한 정보를 제공합니다.

Table 5에도 동일한 패턴이 수치적으로 나타납니다. 슬로프는 100 Elo 포인트당 측정값 변화로 보고됩니다. 캐슬링, 마이너피스 전개, 반복된 곡약장, 초기 퀸약장 모두 명확하게 나타납니다. 600-1600에 걸친 움직임. 1600 이상에서는 경사가 훨씬 작습니다. 이는 더 높은 등급의 플레이어가 깰 수 있다는 생각과 일치합니다. 기본 원칙을 보다 선택적으로, 보다 성공적으로 수행할 수 있습니다. 두 번째 메커니즘은 채도. 일부 습관의 경우 측정항목이 실제 상한선에 도달하거나 적정 수준에 도달합니다. 스팟: 강한 선수들은 이미 습관을 배웠기 때문에, 그럴 이유가 없습니다. 곡선은 지속적으로 급격하게 상승할 것입니다. 더 많은 것이 항상 더 나은 것은 아닙니다. 그 점. 따라서 더 강한 플레이어가 깰 수 있기 때문에 데이터는 두 가지를 모두 평평하게 만듭니다. 구체적인 입장에서 규칙을 정하고 이미 많은 기본 습관에 도달했기 때문입니다. 유용한 작동 범위.

10 상위 컷오프가 중요한 이유

Elo+Chess 대상 사용자는 초급부터 고급까지의 플레이어입니다. 이들을 위해 플레이어의 기본 습관은 평가와 밀접한 관련이 있습니다. 반복되는 초기 조각 이동 방지, 불필요한 초기 퀸 이동 방지, 성 시간이 지나면 루크를 연결하고, 재료를 관리하고, 조각을 느슨하게 두거나 교수형.

등급이 높을수록 동일한 표면 습관이 덜 진단됩니다. 강한 중앙이 닫혀 있기 때문에 플레이어는 캐슬링을 지연시킬 수 있습니다. 여왕을 일찍 이동시키세요. 위치에 구체적인 전술이 포함되어 있거나 동일한 조각을 두 번 이동하기 때문입니다. 템포를 얻거나 구조적 양보를 강요하기 때문입니다. 아직도 원시적인 습관 의미가 있지만 습관과 평가 사이의 관계는 더욱 중요해집니다. 조건부. 이것이 Elo+Chess가 메인 코칭에 상위 컷오프를 사용하는 이유입니다. 주장과 해당 범위 이상의 사용자에게 긴밀한 관계가 있다는 경고가 표시되는 이유 원칙적인 습관과 평가 사이가 무너집니다.

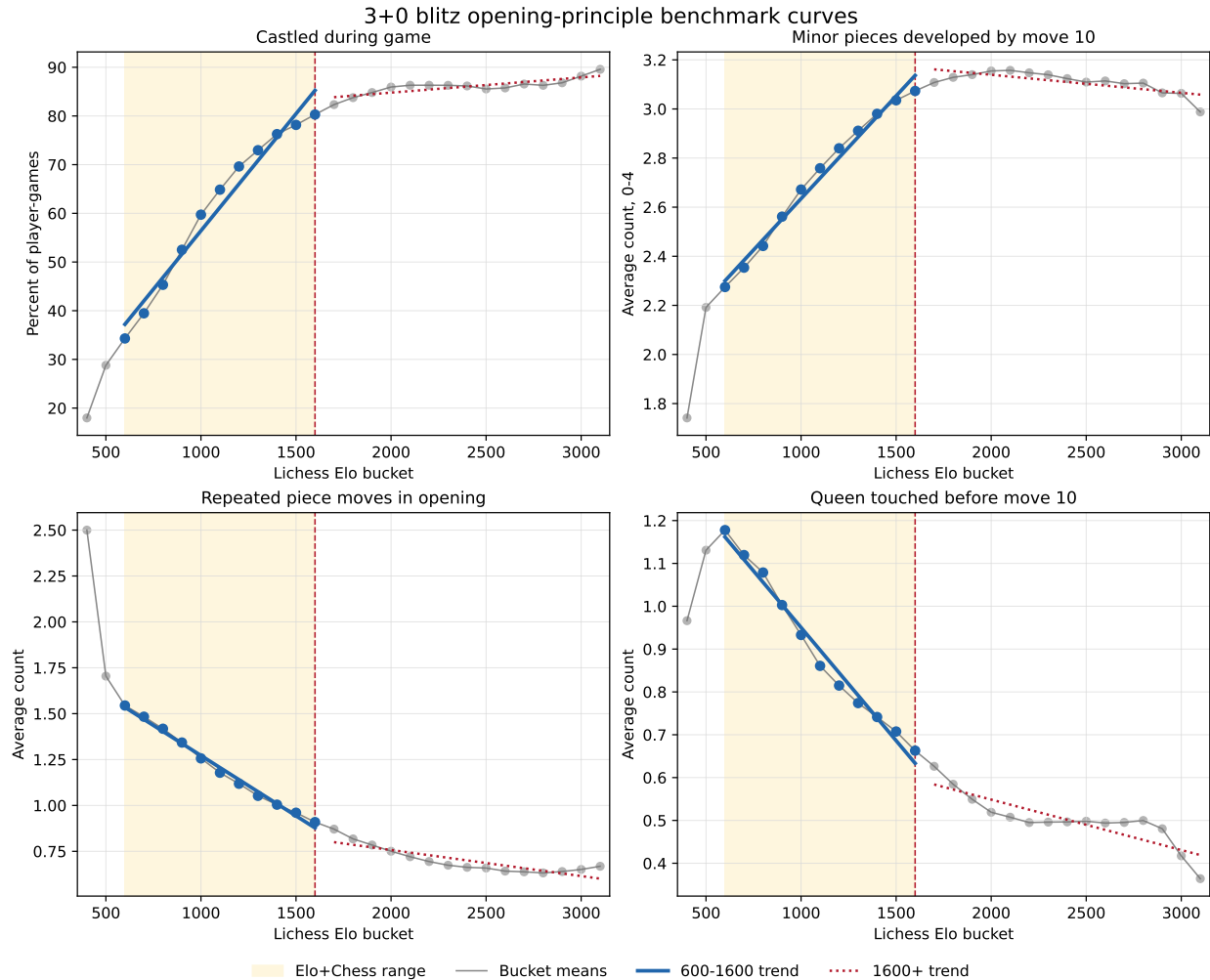


Figure 5: 3+0 블리츠에 대해 선택된 개방 원리 벤치마크 곡선. 는 관계는 Elo+Chess 코칭 범위에서 가장 강력하고 유용합니다. 상위 컷오프 위에서는 여러 관계가 상당히 평탄해졌습니다.

11 양면 전술 지표 및 백분위수 목표

일부 메트릭은 개방 원리 예제와 반대되는 모양을 갖습니다. 포크 이벤트는 본질적으로 양면적이기 때문에 유용한 예입니다. 선수는 동료보다 더 유용한 포크를 찾아서 변환하는 것은 좋은 일을 하고 있지만, 변환된 포크의 인구 비율은 등급에 따라 감소합니다. 상대방은 더 적은 포크 기회를 허용하고 전술적 위협으로부터 방어합니다. 더 적극적으로.

그림 8는 이 효과를 보여줍니다. 둘 다 포크를 변환했습니다. 보고서 플레이어와 상대방이 변환한 포크는 다음과 같이 덜 빈번해집니다. 등급이 상승합니다. 그렇다고 포크 실행이 나쁘다는 의미는 아닙니다. 원시 이벤트를 의미합니다. 비율은 상대 오류 비율과 게임 질감에 의해 부분적으로 제어됩니다. 측정항목의 경우 따라서 Elo+Chess는 내부에서 피어 기록 백분위수를 사용합니다. 글로벌 버킷 평균 추세에만 의존하기보다는 플레이어 자신의 Elo 버킷을 사용합니다.

이러한 백분위수 목표를 구축하기 위해 Elo+Chess는 게임 유형별로 플레이어를 샘플링하고 Elo 버킷은 샘플링된 각 플레이어에 대해 최근 적격 게임을 사용합니다. 는 현재 목표 세트는 600에서 1600까지의 11개 버킷을 사용하며 각 플레이어 수는 200명입니다. 버킷, 각 프로덕션에 대해 플레이어당 최근 적격 게임 최대 50개 게임 유형. 이는 게임 유형당 2,200개의 플레이어 샘플을 제

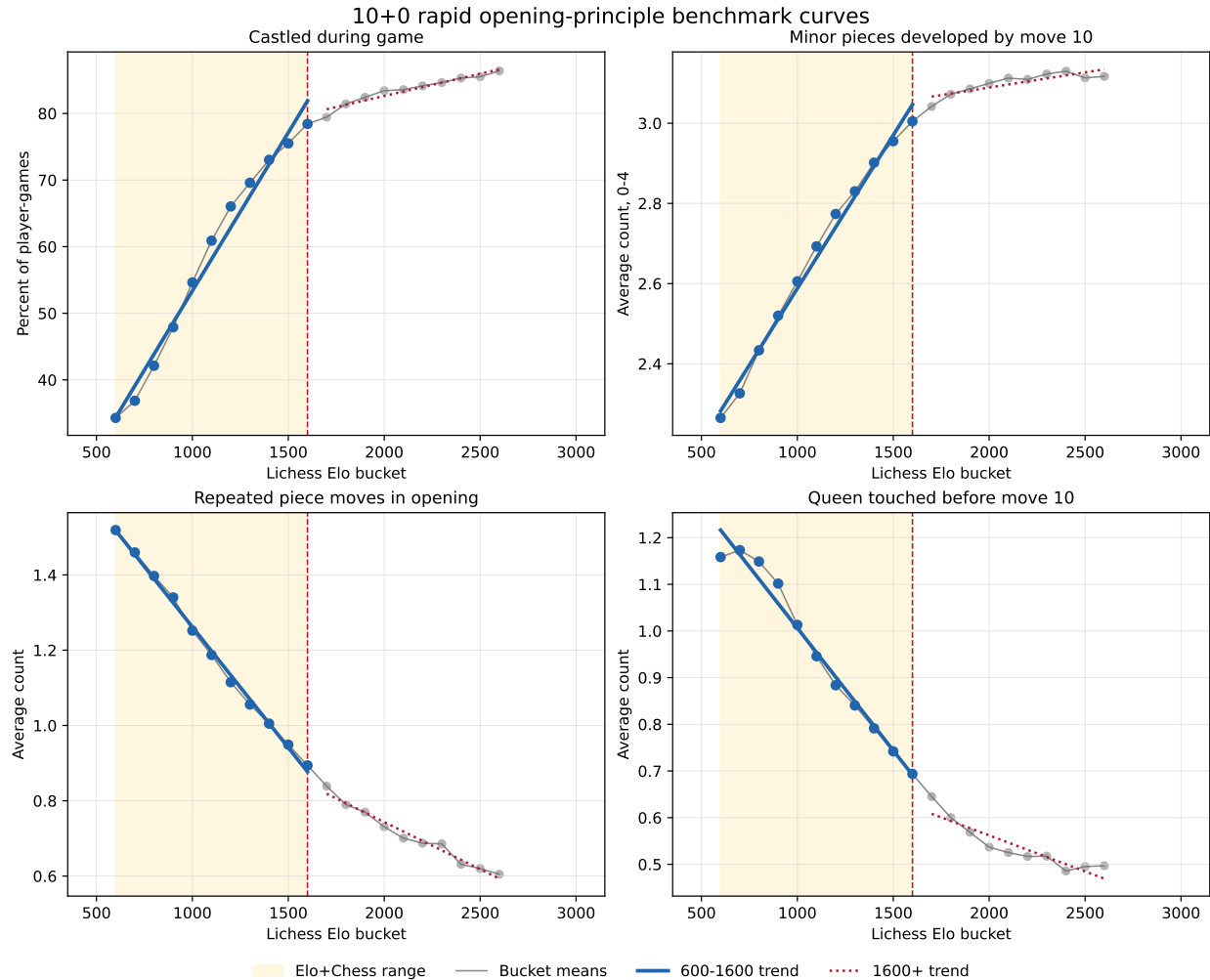


Figure 6: 10+0 래피드 게임에 대해 선택된 개방 원리 벤치마크 곡선. 는 관계는 Elo+Chess 코칭 범위에서 가장 강력하고 유용합니다. 상위 컷오프 위에서는 여러 관계가 상당히 평탄해졌습니다.

공하며 대략적으로 버킷 내 백분위수에 대한 게임 유형당 플레이어-게임 행 110,000개 배포판.

그림 9는 결과 타겟을 보여줍니다. 3+0 블리츠에서 샘플링된 플레이어 변환 포크 비율에 대한 분포입니다. 각 상자는 버킷 평균이 아닌 샘플링된 플레이어 전체의 버킷 내 분포 개인 게임. 보고서 UI에서는 사용자의 가치를 관련 가치와 비교합니다. 사용자의 매핑된 Elo 버킷 상자입니다. 이는 다음과 같은 진술의 기초가 됩니다. 플레이어의 변환된 포크 비율이 동료에 비해 높거나 낮은지 여부 등급에 따라 전체 모집단 사건 발생률이 감소하는 경우에도 동일한 수준입니다.

12 복제 비용

이 방법은 원칙적으로 재현 가능하지만 가볍지는 않습니다. 재건축 벤치마크 곡선에는 다음이 필요합니다.

- 대규모 Lichess 월별 게임 기록 파일에 액세스
- 게임 유형, 등급 버킷, 자격에 따라 계층화된 필터링 규칙;

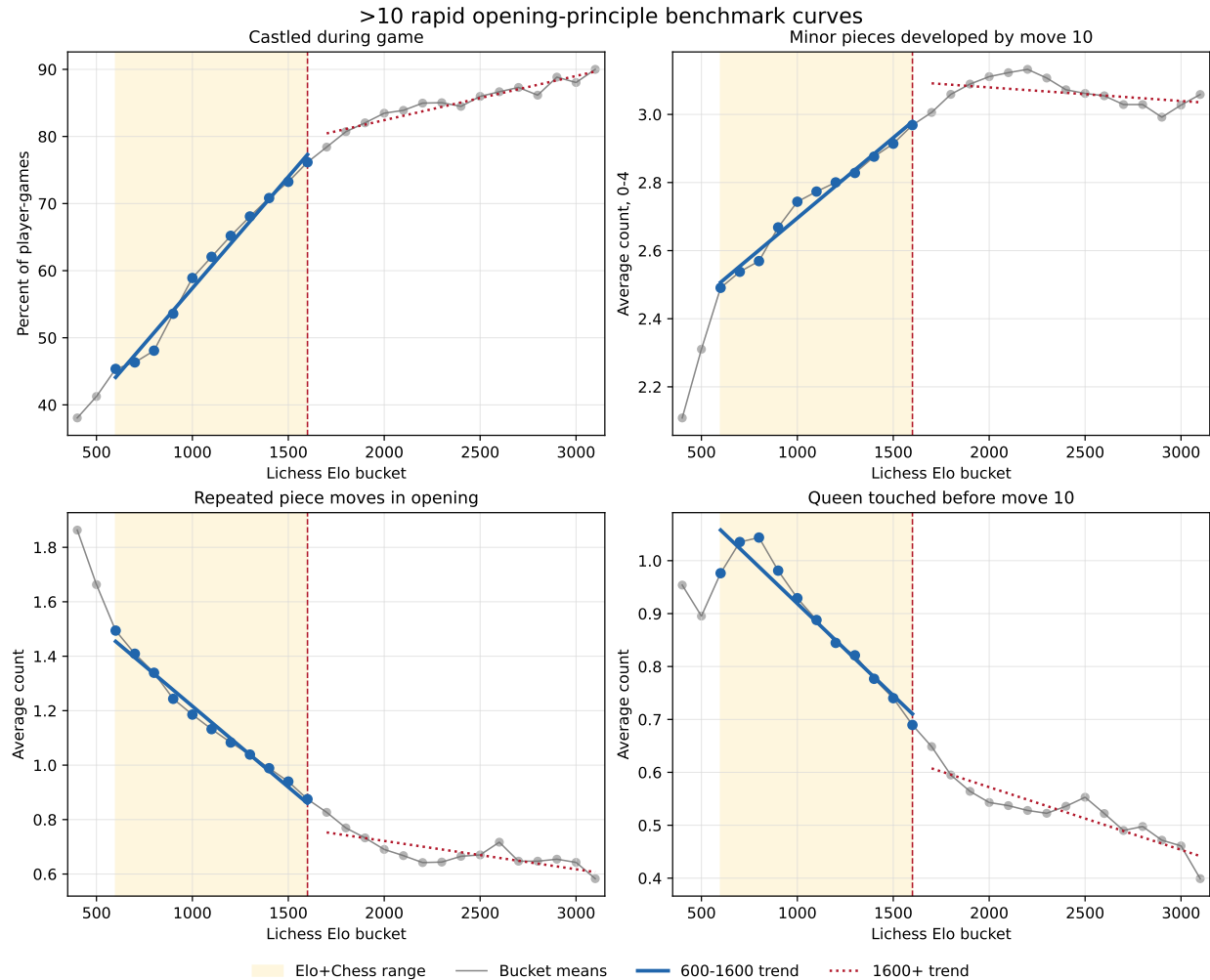


Figure 7: 더 길고 빠른 게임을 위해 선택된 개방 원리 벤치마크 곡선.

- 확장 가능한 법적 이동 파서 및 보드 상태 평가자 수천만 개의 이동 상태;
- 오프닝, 재료, 폰 구조, 킹 안전, 루크에 대한 기능 로직 활동, 전술, 시간 관리 및 게임 결과;
- 이동 상태에서 플레이어-게임 기능으로의 집계
- 모든 보고서 지표에 대한 버킷 수준 요약
- 모든 후보 벤치마크 곡선을 사용하기 전에 수동으로 검토 코칭 보고서.

3+0 블리츠의 경우 이는 거의 50만 게임, 거의 100만 게임을 의미합니다. 플레이어-게임 행, 2,800만 개 이상의 이동 상태 행이 있습니다. 4개에 걸쳐 벤치마크 게임 유형, 현재 벤치마크 아티팩트는 2.55 이상을 포함합니다. 백만 개의 게임, 510만 개 이상의 플레이어-게임 행, 147개 이상의 백만 개의 이동 상태 행. 규모는 가치의 중심이다. 벤치마크: 플레이한 게임에서 실제로 무슨 일이 일어나는지 보고서를 통해 사용자에게 보여줄 수 있습니다. 일반적인 체스 슬로건에만 의존하기보다는 자신의 수준에 가까운 플레이어가

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: 네 가지 개방 원리 예시에 대한 대략적인 버킷 곡선 경사 프로덕션 게임 유형 전반에 걸친 측정항목.

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: 현재 피어 기록 백분위수 목표 세트입니다. 이들은 버킷 내 분포를 추정하는 데 사용되는 플레이어 기록 샘플 백분위수 스타일 측정항목.

13 제한 사항

벤치마크 곡선은 인과관계 증명이 아닌 경험적 설명입니다. 플레이어 단순히 버킷 평균을 일치시키는 것만으로는 등급을 얻지 못합니다. 측정항목도 복잡한 위치에 대한 불완전한 요약: 초기 여왕 이동은 훌륭할 수 있습니다. 한 위치에서는 무모하고 다른 위치에서는 무모합니다. 곡선의 목적은 다음과 같습니다. 반복되는 습관을 파악하고 이를 오래 행동과 비교합니다. 계산이나 코칭 판단을 대체합니다.

곡선은 Lichess 스케일 곡선이기도 합니다. 사용자가 Chess.com 게임을 가져올 때 기록, Elo+Chess는 관련 등급 버킷을 Lichess 벤치마크에 매핑합니다. 별도의 크로스 플랫폼 매핑 모델을 사용하여 확장합니다. 해당 매핑 레이어는 Lichess-to-Chess.com에 대한 동반 Elo+Chess 연구 노트에 설명되어 있습니다. 평가 변환: [Estimating Cross-Platform Chess Rating Mappings with 모달 회귀](#). 이 매핑은 백분위수를 유지하는 것으로 읽어서는 안 됩니다. 순위. 두 등급 사이의 추정 등급 동등성을 유지합니다. 시스템; 백분위수 순위는 활성 플레이어가 플랫폼에 따라 달라지기 때문에 유지됩니다. 분포도가 다릅니다.

14 생산 보고서 그래픽

생산 보고서는 벤치마크 아티팩트를 메트릭 수준 패널로 전환합니다. 전체 빌드 파이프라인을 몰라도 읽을 수 있습니다. 피규어 10 and 11 라이브 Elo+Chess 보고서 디자인의 세 가지 예를 보여

Two-sided converted-fork rates by Elo bucket

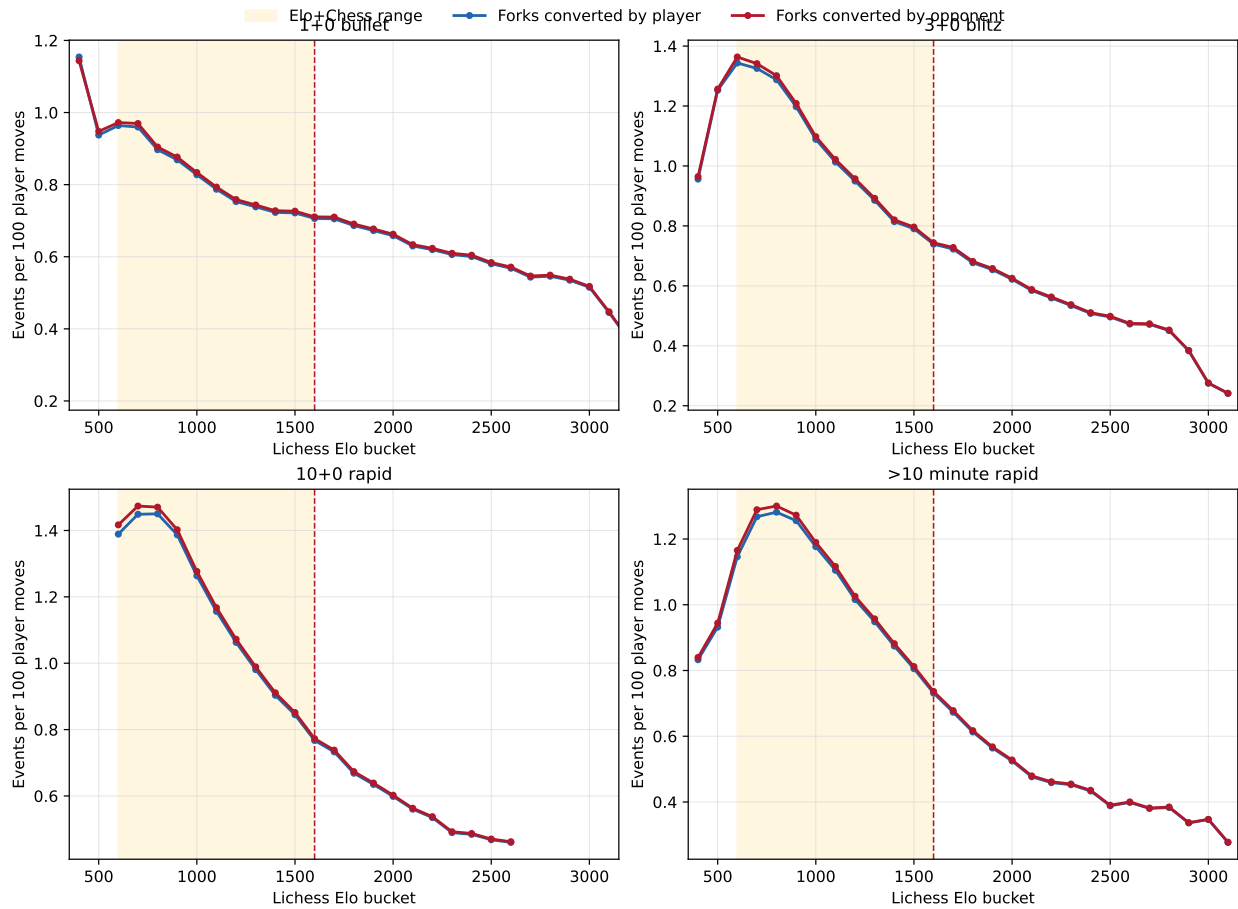


Figure 8: Elo 버킷에 의한 양면 변환 포크 비율. 변환된 포크 이벤트 양쪽 모두 더 잘 방어하고 더 적은 수의 깨끗한 포크를 방어하기 때문에 더 높은 등급으로 떨어집니다. 전환이 가능합니다.

줍니다.

Figure 10의 개방 원리 패널 일반적인 벤치마크 곡선 프레젠테이션을 사용합니다. 각 카드는 다음으로 시작됩니다. 측정항목군과 측정항목 이름을 입력한 다음 간단한 해석 규칙을 제공합니다. 오른쪽 상단의 큰 등급은 의도적으로 원시 등급과 분리되어 있습니다. 측정항목 값: '이 습관 등급은 어떻게 매기는가?'라는 사용자 질문에 답합니다. 벤치마크 대비?" 반면 아래 행은 측정된 값을 유지합니다. 4개의 요약 행에는 수명, 최근 10개, 승리 및 손실이 표시됩니다. 이는 단순한 장식이 아닌 패널 진단. Lifetime과 Last 10이 일치하는 경우 습관은 아마도 안정적 일 것입니다. 승패가 다른 경우 측정항목은 다음과 같습니다. 단순한 습관이 아닌 결과에 의존하거나 게임 상태에 묶여 있습니다.

차트 영역은 절제된 어두운 배경, 금색 추세선 및 컬러를 사용합니다. 네 조각에 대한 마커입니다. 추세 주석은 다음 사항을 명시적으로 나타냅니다. 벤치마크 관계는 Elo와 함께 상승하거나 하락합니다. 이것이 중요한 이유는 작은 부분 개발과 같은 일부 측정 항목은 높을수록 좋습니다. 초기 여왕 운동과 같은 다른 것들은 낮을수록 좋습니다. 그러므로 카드는 사용자에게 경사면만으로 방향을 추론하도록 요청하지 않습니다. 결승전 '이것이 무엇을 의미하는가'와 '코칭 팁' 라인은 벤치마크 결과를 번역합니다. 경험적 맥락을 보존하면서 구체적인 체스 교육으로 전환합니다.

그림 11는 백분위수 스타일을 보여줍니다. 원시 비율이 잘 표현되지 않는 측정항목에 사용되

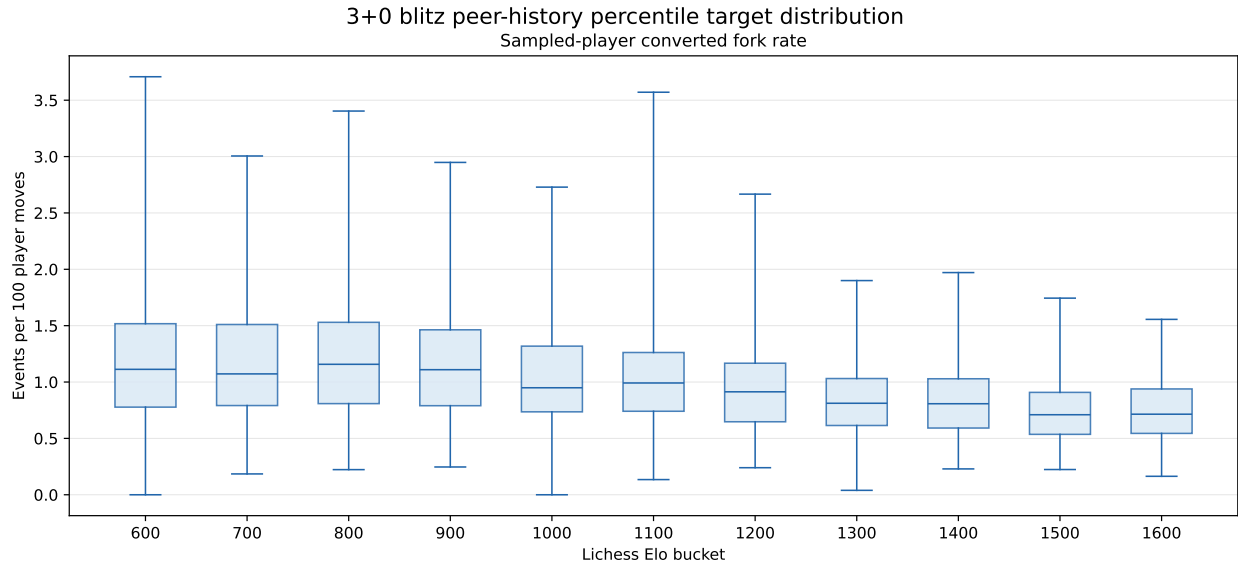


Figure 9: 예시 3+0 샘플링된 플레이어에 대한 블리츠 피어 히스토리 목표 분포 변환된 포크 속도. 백분위수는 Elo 버킷 내에서 계산되므로 보고서는 대체로 유사한 전술적 기회에 직면한 동료와 플레이어를 비교합니다. 요금.

는 프레젠테이션 단일 글로벌 트렌드. 승리 포크는 전술적 예입니다. 인구 비율 변환된 포크의 비율은 플레이어 기술과 상대 오류율에 따라 변경됩니다. 는 따라서 카드는 사용자의 가치를 동료 플레이어 기록과 비교합니다. 매핑된 Elo 버킷. 상자 그림에는 다음을 포함한 동료 분포가 표시됩니다. 사분위간 범위, 중앙값 및 수염. 사용자의 마커가 척도가 동일하고 원시 값과 백분위수 순위가 모두 표시되어 있습니다. 인접한 비교 버킷은 음소거 형식으로 계속 표시되므로 사용자는 피어 버킷은 범용 규모가 아닌 로컬입니다.

이 설계를 통해 백분위수 측정항목을 감사할 수 있습니다. 사용자는 원시 값을 보고, 백분위수, 비교에 사용된 피어 버킷, 그리고 번역된 Chess.com-Lichess 버킷 매핑. 동일한 수명, 지난 10회, 승리 및 손실 행이 유지되므로 전술적 백분위수 점수를 계속 확인할 수 있습니다. 최신 효과 및 결과 의존성을 위해.

15 결론

Elo+Chess 벤치마크 보고서는 대규모 이동 상태 계산을 통해 작성되었습니다. 계층화된 Lichess 샘플. 각 측정항목은 구체적인 보드 상태 세트로 시작됩니다. 이동 상태 변수는 플레이어-게임 기능이 된 다음 평균화됩니다. Elo 버킷을 사용하여 경험적 벤치마크 곡선을 형성합니다. 개방 원리에는 이 접근 방식이 왜 유용한지를 보여줍니다. 간단한 체스 습관은 강력하고 목표 범위의 등급과 가시적인 관계가 있는 반면, 관계는 종종 상위 컷오프 이상으로 평탄화되거나 조건부로 변합니다. 이것이 Elo+Chess가 등급에 코칭 주장을 집중하는 핵심 이유입니다. 데이터가 이를 가장 명확하게 뒷받침하는 범위입니다.

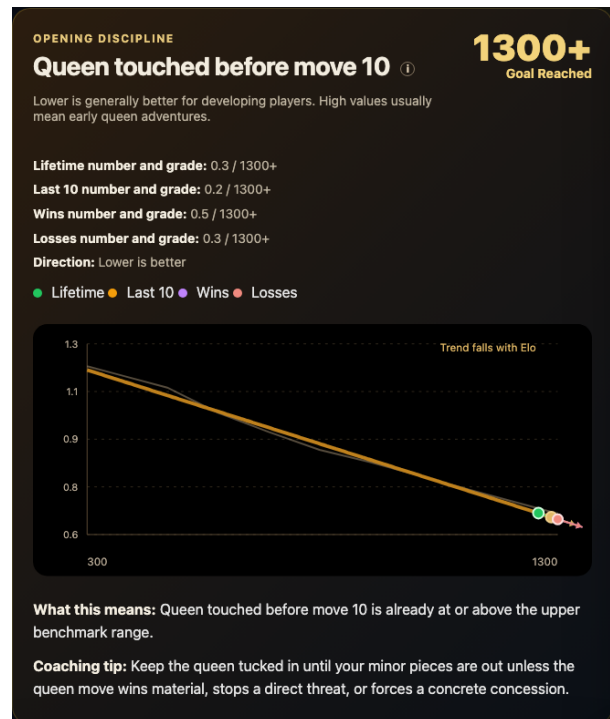
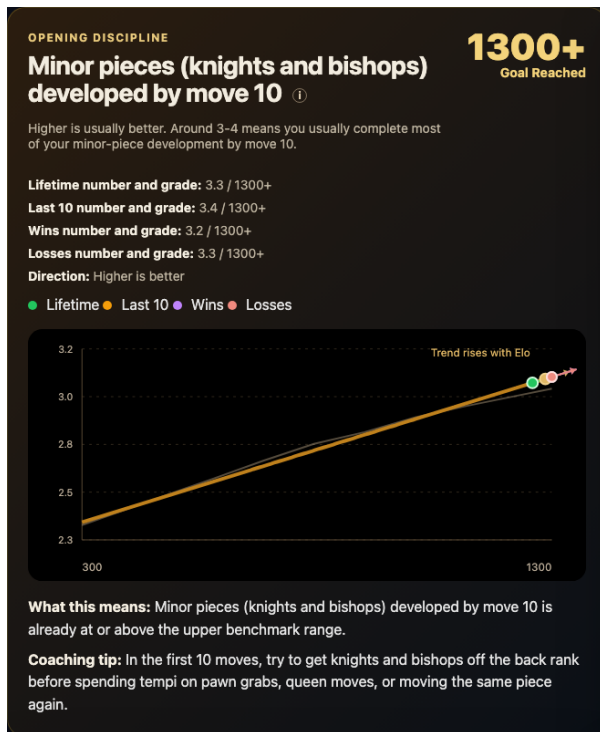


Figure 10: 생산 개방 원리 패널. 동일한 시각적 문법이 높을수록 더 나은 측정법, 작은 부분 개발, 낮을수록 더 나은 측정법, 초기 여왕 운동. 두 패널 모두 수명, 최근 10개, 승리 및 손실을 표시합니다. 따라서 보고서는 안정적인 습관을 최근 또는 결과에 따른 습관과 구분할 수 있습니다. 패턴.

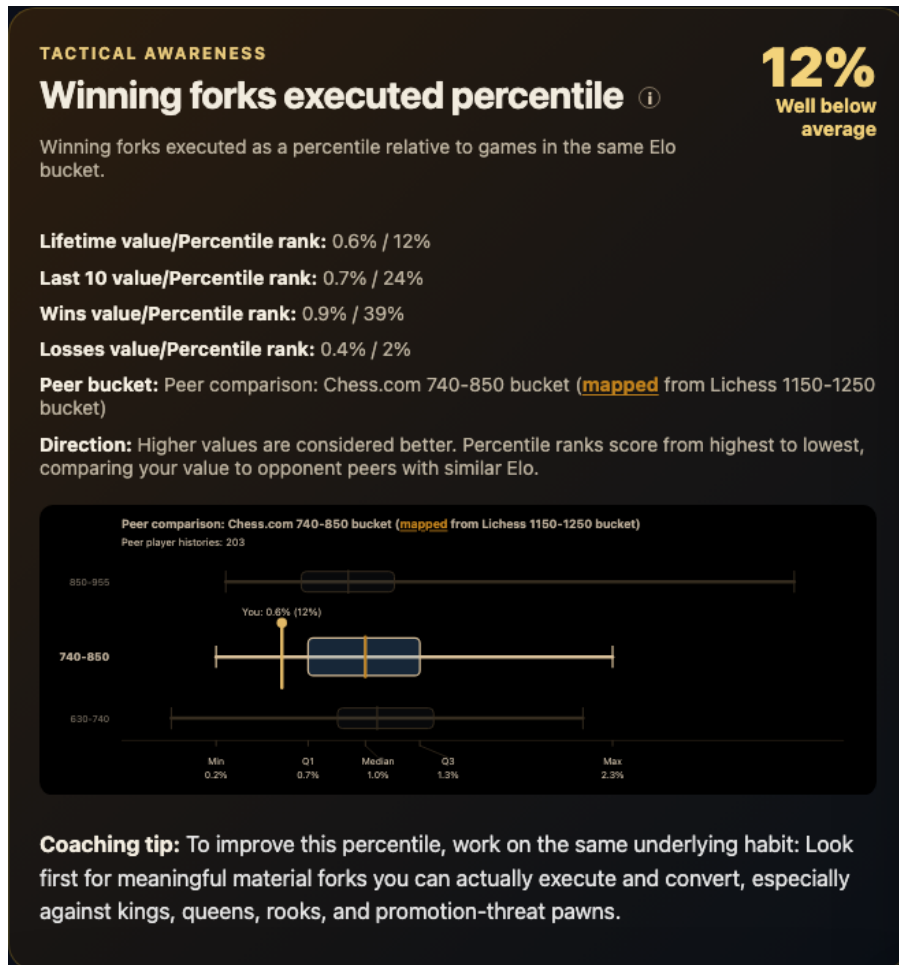


Figure 11: 승리한 포크를 위한 생산 백분위수 패널이 실행되었습니다. 상자 그림 보고서에 사용된 동일 버킷 피어 분포를 보여주고, 골드는 마커는 사용자의 원래 값과 백분위수 순위를 표시합니다.