

## Fantasy Premier League. Предсказатель

В [первой заметке](#) цикла я рассказал, что ряд менеджеров Fantasy Premier League демонстрирует высокие результаты на протяжении многих сезонов. Я выделил менеджеров, входящих в TOP-10k по итогам пяти сезонов 2017/2018–2021/22 (*элиты*), а также группу сравнения – аккаунты, отобранные случайным образом (*поляну*). [Далее](#) я показал, какие возможности для извлечения данных с сайта Fantasy Premier League предоставляет API и Excel Power Query. В [третьей заметке](#) я исследовал различия между элитой и поляной в выборе команды к первому туру. В [четвертой заметке](#) показал отличия в трансферном поведении. В [пятой заметке](#) рассказал об игре фишками, в [шестой](#) – о выборе капитана, в [седьмой](#) – об автозаменах. Сейчас на примере сезона 2022/23 покажу, как фактические результаты коррелируют с прогнозами.

	xMins	GW38	Total	PPM	Elite%
 Kane FW 11.5	82	5.7	5.7	0.50	3.9%
 Jesus FW 8.1	77	5.5	5.5	0.68	0.9%
 Watkins FW 7.3	77	5.0	5.0	0.68	11.0%
 Haaland FW 12.4	59	4.4	4.4	0.36	98.9%
 Wilson FW 7.3	66	4.3	4.3	0.59	10.0%
 Maupay FW 5.6	57	4.2	4.2	0.74	0.0%
 Edouard FW 5.0	59	3.8	3.8	0.76	0.3%
 Isak FW 7.0	67	3.8	3.8	0.54	52.3%

Рис. 1. Прогнозные очки форвардов на GW38 сезона 2022/23 с сайта fplreview

Уже несколько лет я использую сервис [fplreview](#). Он предлагает прогноз очков на ближайшие 8 туров. Поскольку результаты в FPL обладают очень большой дисперсией, факт от прогноза в одном туре для одного игрока могут отличаться, как небо и земля. И, естественно, никой выстрел спрогнозировать невозможно. Однако в среднем прогноз fplreview является лучшим предиктором из всех мне известных. С доказательства этого мы и начнем.

### *Оценка предсказательной силы fplreview*

В течение сезона 2022/23 я собирал прогноз fplreview накануне дедлайна каждой недели, а затем сравнивал этот прогноз с итогами тура. Прогноз на GW19 я не успел получить, и эту неделю исключил из анализа. Корреляция между прогнозом и фактом по каждому игроку за сезон составила 0,97. Корреляция для каждого тура колеблется в диапазоне  $0,60 \pm 0,05$ .

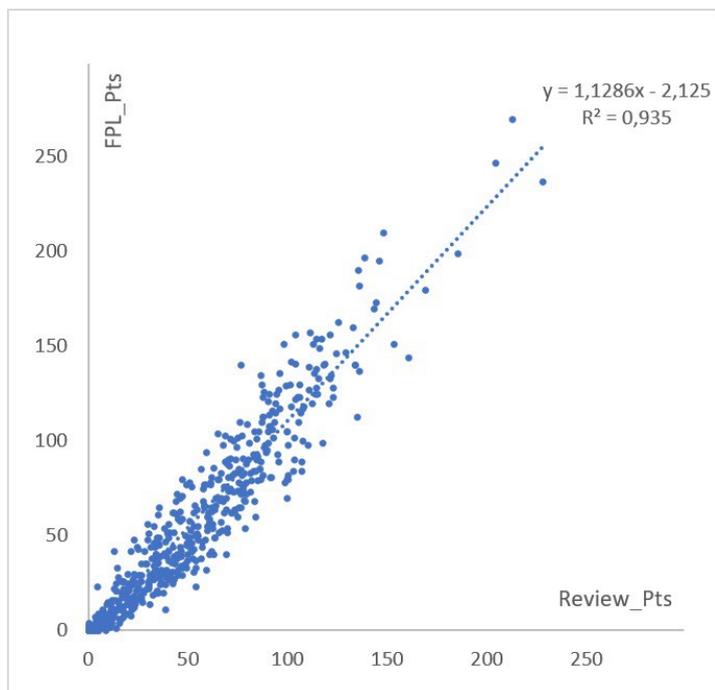


Рис. 2. Корреляция между прогнозом (Review\_Pts) и фактом (FPL\_Pts) для каждого игрока за сезон 2022/23

Уравнение регрессионной кривой показывает, что фактически набранные очки закономерно превосходят прогноз с коэффициентом регрессии 1,13. Т.е. в среднем прогноз занижен, и авторы модели могли бы внести соответствующие коррективы. На рис. 2 также видно, что среди игроков, есть как «везунчики», так и «неудачники». Для них фактические показатели значительно отличаются от прогноза. Рассмотрим лидеров сезона:

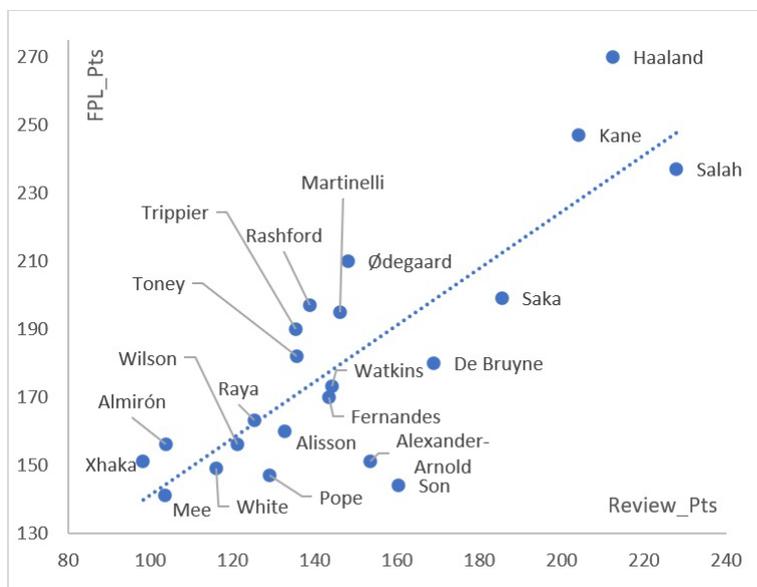


Рис. 3. Прогноз и факт для лидеров сезона 2022/23

Сильно недобрали до прогноза: Son, TAA, Pope, De Bruyne, Saka. Набрали больше, чем предсказывал fplreview: Haaland, Ødegaard, Rashford, Trippier, Kane, Martinelli.

Менеджеры по ходу сезона часто ориентируются на популярность игроков. Так вот, корреляция между загрузками и результатом для отдельных туров колеблется в диапазоне  $0,33 \pm 0,18$ , а в целом за сезон составляет 0,66. Еще один популярный предиктор – очки, набранные игроком. Корреляция между Pts, набранными во всех предыдущих турах, и очками в текущем туре для отдельных туров колеблется в диапазоне  $0,50 \pm 0,05$ . Т.е., значительно лучше, чем загрузки, но уступает прогнозу fplreview.

Итак, fplreview – статистически лучший предиктор результатов.

### Мини-лига

Я участвую в двух мини-лигах и по итогам сезона 2022/23 для участников этих лиг подсчитал прогноз и факт:

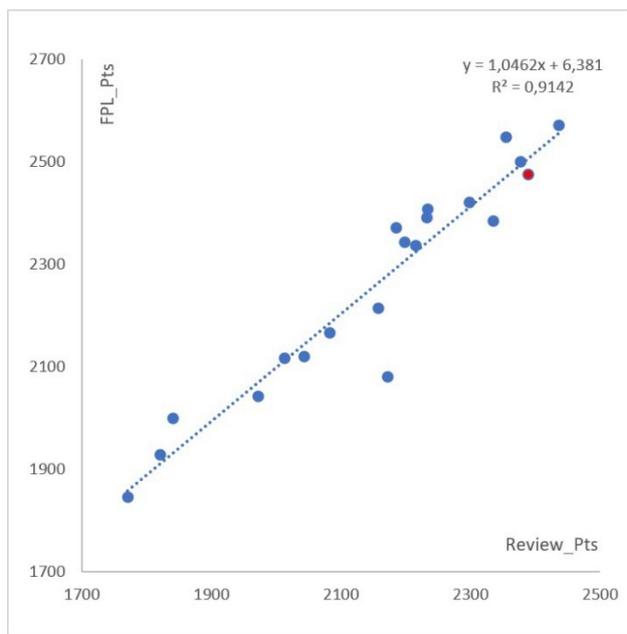


Рис. 4. Мини-лиги; мой результат выделен

Видно, что я занял второе место по прогнозным результатам, но только четвертое по факту. Корреляция между прогнозом и фактом составила фантастические 95,6%. Для наших лиг коэффициент регрессии меньше, чем глобальный и составляет 1,046. Но опять же фактические результаты закономерно превосходят прогнозные.

### Элита и поляна

Визуально отличия в поведении элиты и поляны не заметны:

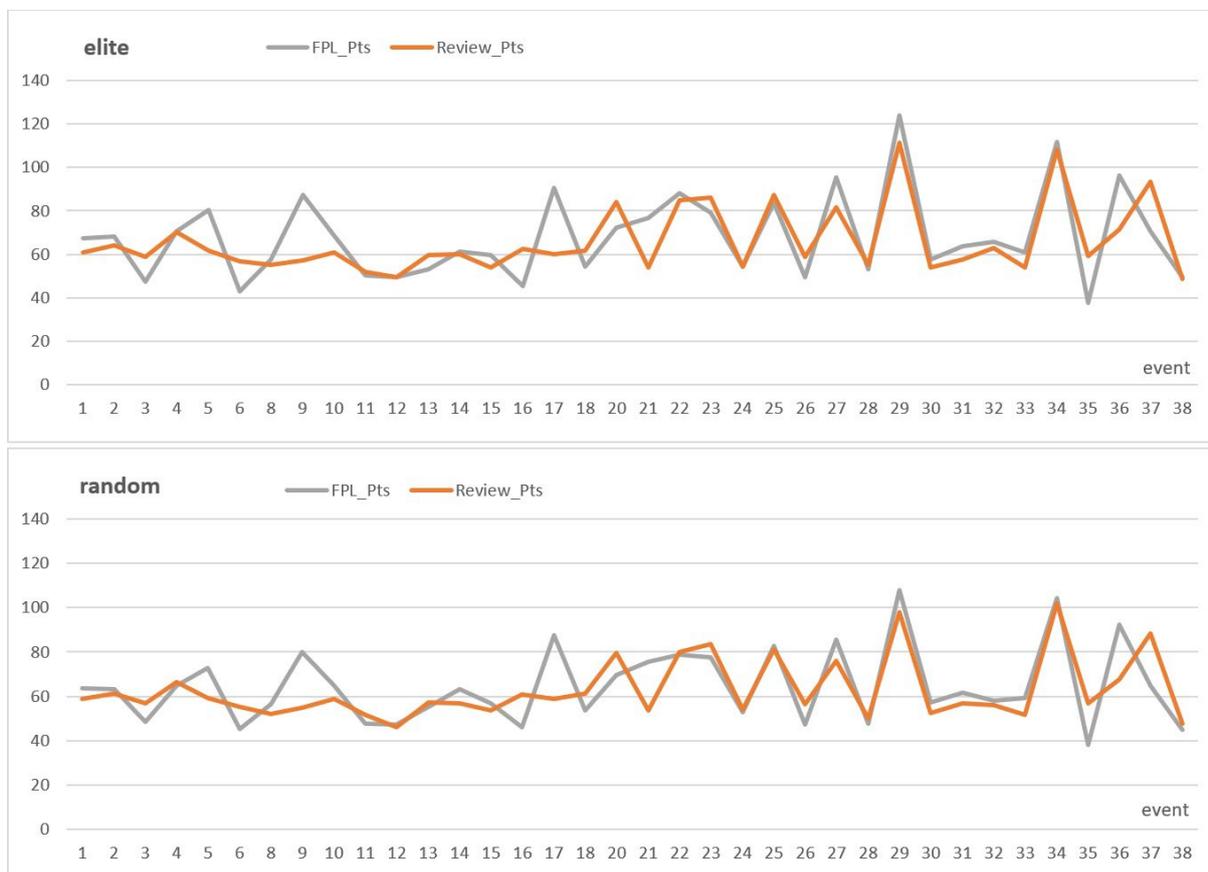


Рис. 5. Колебания прогноза и факта отдельно по элите и поляне по турам сезона 2022/23

В большинстве туров прогноз и факт совпадают. Но есть туры, в которых прогноз выше или ниже факта, приблизительно одинаково для элиты и поляны.

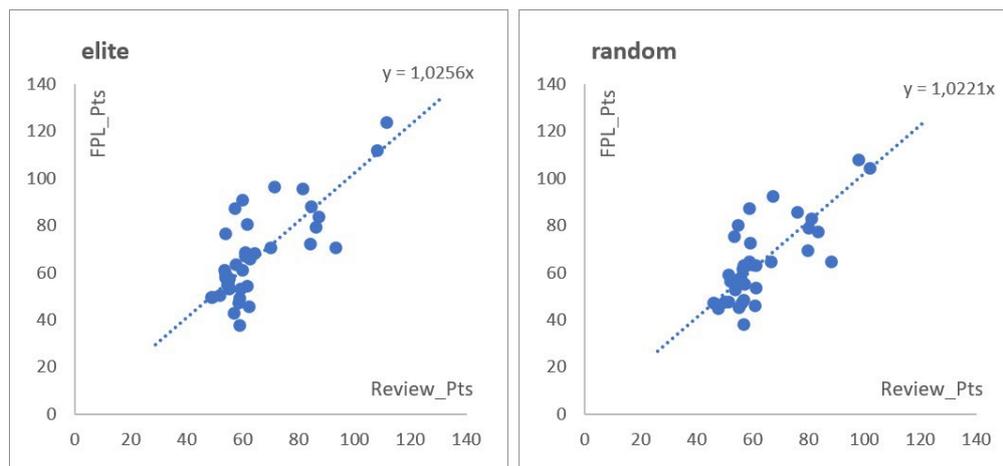


Рис. 6. Корреляция прогноза и факта отдельно для элиты и поляны; точка – один тур

Различия между группами становятся очевидны, когда мы оценим дельту средних значений прогнозных и фактических очков по турам:

$$\Delta\text{Review\_Pts} = \text{Review\_Pts}_{\text{elite}} - \text{Review\_Pts}_{\text{random}}$$

$$\Delta\text{FPL\_Pts} = \text{FPL\_Pts}_{\text{elite}} - \text{FPL\_Pts}_{\text{random}}$$

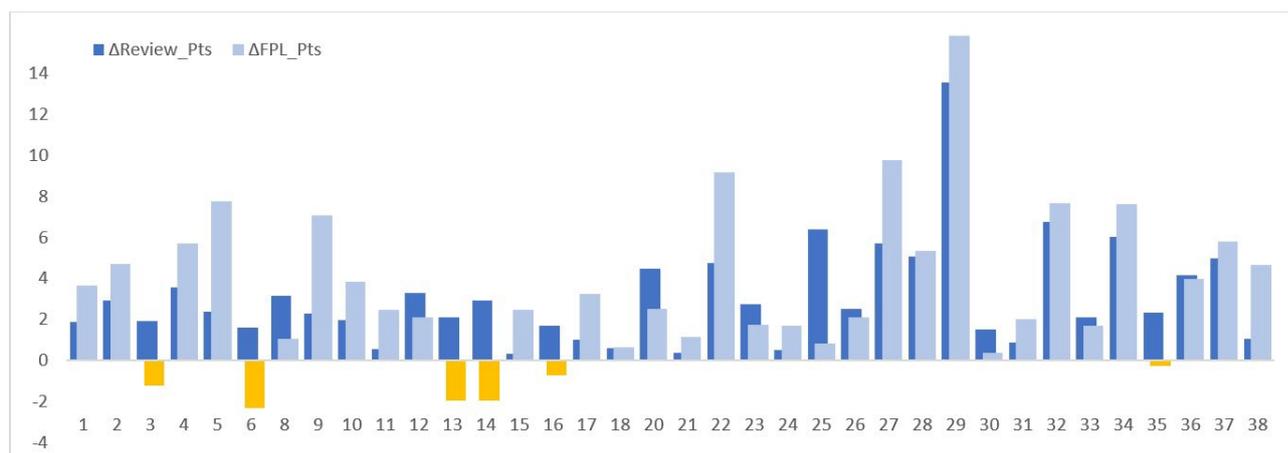


Рис. 7. Дельта для элиты и поляны: 1) между прогнозами и 2) между фактическими результатами

Темно-синие столбики показывают разницу по турам между прогнозом для элиты и поляны. Все они больше нуля! Светло-синие столбики (с инверсией желтым) показывают дельту между фактическими результатами элиты и поляны. В шести турах поляне в среднем удалось превзойти элиту.

По всем турам средняя разница прогнозов составила 3,1 очка за тур, а средняя разница результатов – 3,3 очка за тур.

Самой «невозучей» для элиты была BDGW25. На эту неделю пришелся финал Кубка лиги, в котором играли MUN и NEW. Это привело к переносу двух матчей. С другой стороны, организаторы поставили вторые матчи Арсеналу и Ливерпулю. Элита загрузила игроков LIV, которые в этот момент сезона были не очень популярными, но их выступление было весьма посредственным.

И, наоборот, одной из самых «везучих» для элиты была DGW22, в которой элита сделала тройным капитаном Рэшфорда, и он оправдал ожидания, набрав отличные 20 очков!

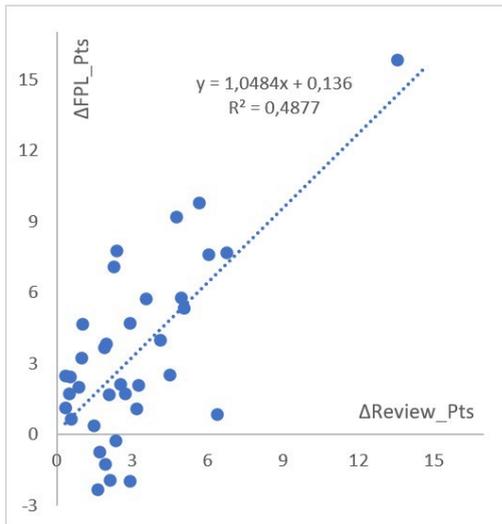


Рис. 8. Корреляция различий в прогнозе и факте

Здесь по оси абсцисс для каждого тура оценивается разница прогнозов  $\text{Review\_Pts}_{\text{elite}} - \text{Review\_Pts}_{\text{random}}$ . А по оси ординат – разница результатов  $\text{FPL\_Pts}_{\text{elite}} - \text{FPL\_Pts}_{\text{random}}$ . Коэффициент корреляции составляет приличные 70%.

Корреляцию между прогнозом и фактом можно отразить для каждого менеджера отдельной точкой.

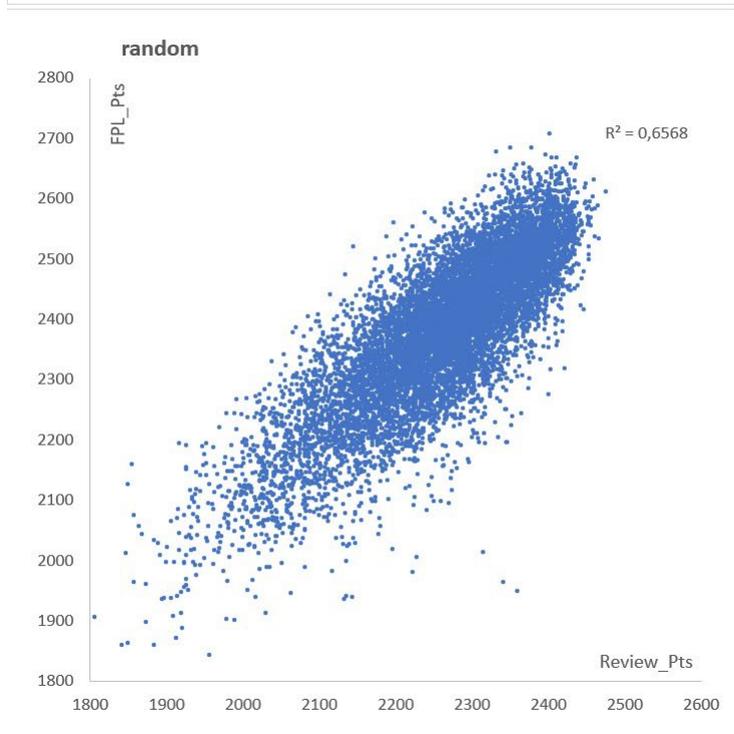
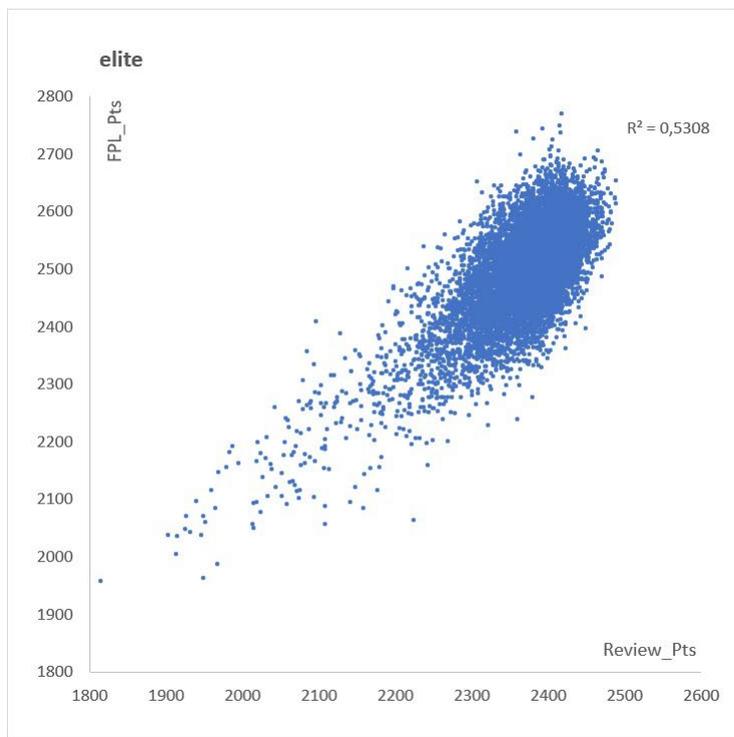


Рис. 9. Корреляция между прогнозом и фактом для каждого менеджера отдельно

Для элиты наблюдается смещение вправо и вверх, т.е., к **большим** значениям, как прогнозных, так и фактических результатов. Значения для элиты собраны в меньшей области, т.е., более концентрированы. Коэффициент корреляции между прогнозом и фактом у элиты несколько меньше, чем у поляны (73% против 81%). Могу предположить, что элита чаще идет против прогноза, вынося более точное суждение, чем модель fplreview. Но это лишь догадка...

Меня интересовал вопрос, использует ли элита прогноз fplreview чаще, чем поляна. Анализ данных не дал ответа. Похоже, что анализ больших данных не поможет в решении вопроса. Нужны интервью с игроками, чтобы пролить на это свет.

Являются ли элитные игроки более (или менее) «везучими», чем поляна? Чтобы ответить на этот вопрос для каждого аккаунта я подсчитал разницу между фактическими результатами за сезон и прогнозом fplreview. Распределение числа игроков по этой разнице для элиты и поляны имеет одинаковое среднее и форму:

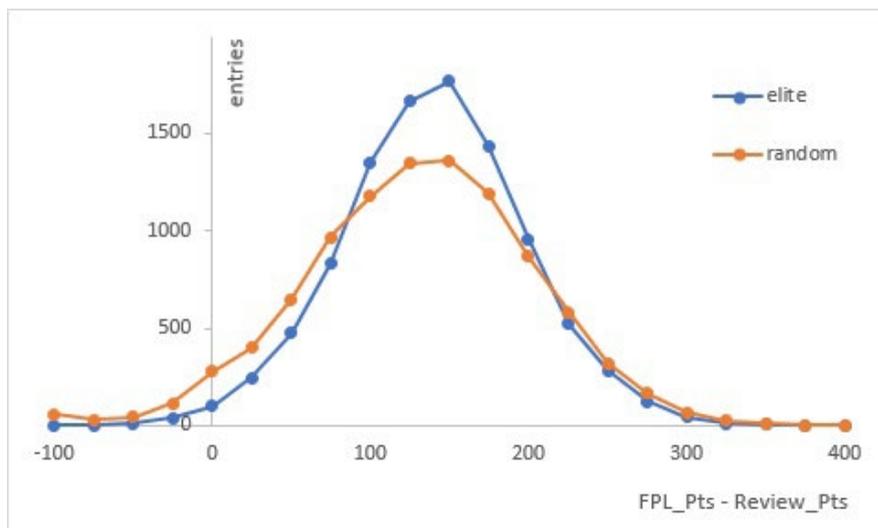


Рис. 10. Распределение числа игроков отдельно для элиты и поляны по разнице FPL\_Pts – Review\_Pts

Если под везением понимать разницу между фактом и прогнозом, то различий между элитой и поляной практически нет. А более высокие результаты элиты проистекают из более высоких прогнозных значений. Формируя состав, основанный на прогнозе, элита последовательно добивается более высоких результатов.

#### TOP-10k сезона 2023/23

Я так же сравнил прогноз и факт элиты и TOP-10k по итогам сезона 2022/23. Ясно, что средние результаты группы TOP-10k выше, чем у элиты, а вот, выше ли прогноз? Или успех TOP-10k связан с большим превышением факта над прогнозом?

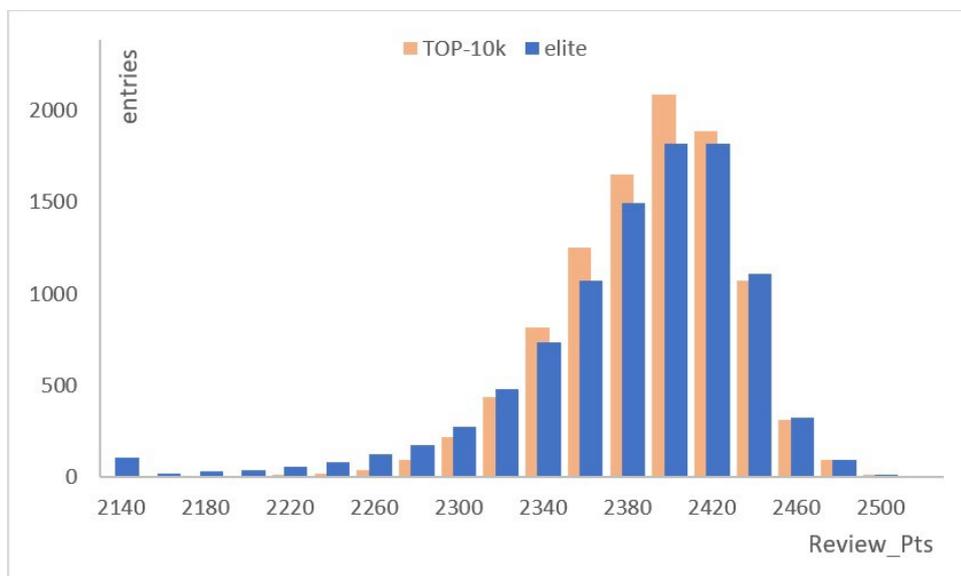


Рис. 11. Распределение прогнозных очков для элиты и TOP-10K

Медианное значение для обеих групп совпадают: 2384,3 для TOP-10k и 2382,9 – для элиты. Да и сами распределения близки. Поскольку даже среди элиты есть менеджеры, которые по ходу сезона теряют интерес к игре, мы видим относительно много небольших значений в диапазоне до 2300 прогнозных очков. В диапазоне 2340–2420 группа TOP-10k превосходит элиту. А для высоких прогнозных значений (2440 и более) элита чуть вырывается вперед.

Но еще интереснее посмотреть на *распределение* превышения фактических очков над прогнозными.



Рис. 12. Распределение превышения фактических очков над прогнозными

Картина радикально отличается от предыдущего сравнения элиты и поляны. Позволю предположить, что превосходство группы TOP-10K над элитой обязано «везению», как мы его определили выше.

### Выводы

Сервис fplreview предлагает наилучший предиктор результатов FPL. В целом по сезону 2022/23 для игроков PL корреляция прогнозных и фактических очков составила 97%. В каждом отдельном туре по всем игрокам PL корреляция находится в диапазоне  $60 \pm 5\%$ .

В каждом туре среднее прогнозируемое количество очков элиты превышало аналогичный показатель для поляны на 3,1 очко. При этом в 6 турах фактические результаты поляны были лучше. Видимо, это были «невезучие» туры.

Анализ данных не позволил выяснить, использует ли элита прогноз fplreview чаще, чем поляна.

По прогнозным результатам элита и группа TOP-10k почти не различаются, а вот по фактическим отличия значительные. Похоже, такую разницу следует отнести на «везение».