**УДК 311.41**

**Иван Сергеевич Кочегаров, студент,**

**Балтийский государственный технический университет имени Д.Ф. Устинова, Россия, г.Санкт-Петербург**

**Болдырева Мария Владимировна, студентка,**

**Балтийский государственный технический университет имени Д.Ф. Устинова, Россия, г.Санкт-Петербург**

**АНАЛИЗ СТАТИСТИЧЕСКИХ ДАННЫХ ПО ИГРЕ В ЗАЩИТЕ В NBA**

***Аннотация****: Данная статья посвящена анализу статистических данных по противодействию нападению в национальной баскетбольной ассоциации, что позволяет точнее оценить влияние на игру каждого игрока и точнее прогнозировать предполагаемые результаты встреч. Для оценки показателей проводится статистическое исследование методом регрессионного анализа.*

***Ключевые слова:*** *Анализ данных, баскетбол, статистический анализ.*

***Abstrac****t: This article is devoted to the analysis of statistical data on counteraction to the attack in the national basketball Association, which allows us to accurately assess the impact on the game of each player and more accurately predict the expected results of the meetings. To assess the indicators, a statistical study is carried out by regression analysis.*

***Key words:*** *data Analysis, basketball, statistical analysis.*

Первым делом необходимо определить, какие же показатели влияют на оценку игры в защите. За основу было взято исследование студентов университета Гарварда «CHARACTERIZING THE SPATIAL STRUCTURE OF DEFENSIVE SKILL IN PROFESSIONAL BASKETBALL» [1]. В данном исследовании предлагается взять за основу такие показатели, как Volume score (число бросков, встреченных защитником), Disruption score (уровень, до которого защитник способен уменьшить эффективность попадания опекаемого игрока), Shot against (взвешенное средние число бросков против данного защитника за 100 владений), counterpoints (взвешенное среднее числа очков, набранных против данного защитника за 100 владений). Данное исследование было проведено в 2014 году и не опиралось на введенную в сезоне 2017-2018 статистику Национальной Баскетбольной Ассоциации по продвинутым показателям (которая позволяет воспользоваться статистикой по сопротивлению броскам оппонентов, отсечениям в борьбе за подбор и касаниям мяча в защите). В результате целью нашего анализа являлось выявление наиболее эффективных в защите игроков именно по продвинутым показателям.

Таблица 1 - этапы анализа

|  |  |  |
| --- | --- | --- |
| **№** | **Название Этапа** | **Результат** |
| 1 | Сбор статистической информации | База данных |
| 2 | Статистическая сводка и обработка информации | Показатели для анализа |
| 3 | Анализ результатов | Выводы |

После определения показателей необходимо собрать данные. Наиболее подробная база данных в открытом доступе, содержащая необходимую для анализа информацию, находится на веб-ресурсе stats.nba.com [2]. Для сбора данных можно воспользоваться API данного сервера. Документация для обращения к API была взята c open-source проекта «nba\_py», размещенного на сайте github.com [3]. В документации не указаны методы для получения статистики по продвинутым показателям, поэтому ссылки для обращения были установлены экспериментально и дополнены. В результате были получены методы leaguehustlestatsteam и leaguehustlestatsplayer, которые возвращают значения показателей бросков с сопротивлением (градация по 2х и 3х очковым), касаний мяча в защите и отсечений при подборах. После определения необходимых методов, с помощью библиотеки Requests для языка Python был написан модуль для сбора данных и сохранения их в формате Pandas DataFrame. После этого были отобраны игроки, сыгравшие не менее 30 игр в сезоне, а также игравшие в среднем не менее 20 минут за игру. Это сделано для того, чтобы точнее оценивать влияние в конкурентной игре против лучших игроков (игроки получающие меньше времени, как правило, выходят против запасных соперника и сами являются запасными). Таким образом была сформирована выборка для анализа. После сбора данных обратимся к регрессионному анализу. Для получения корректной суммы защитных продвинутых показателей максимальное значение по каждому из них было принято за единицу. Из нашего предыдущего исследования [4] были взяты корреляционные коэффициенты продвинутых показателей. После чего значения умножались на корреляционный коэффициент и складывались. Затем для аппроксимации данных необходимо провести прямую с помощью метода наименьших квадратов.

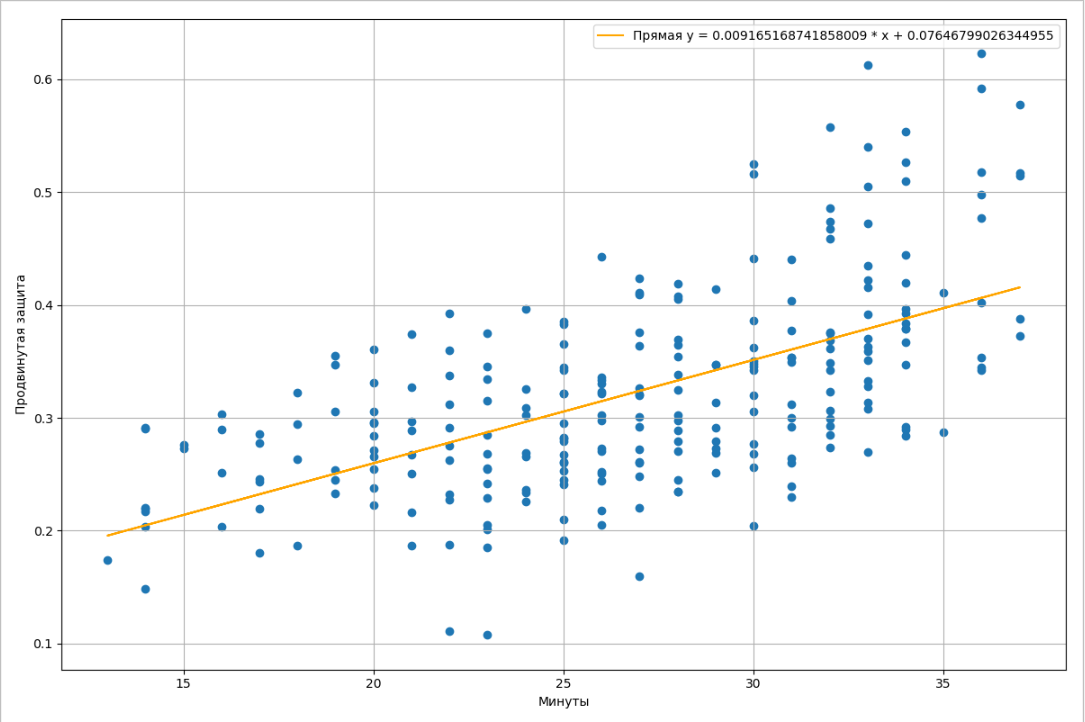


Рисунок 1 – Продвинутая защита от количества минут

С помощью метода linalg.lstsq библиотеки numpy получаем уравнение прямой **y=0.009165168741858009\*x+0.07646799026344955**. В результате получается, что чем выше от линии находится игрок, тем лучше он защищается (рисунок 1).

Теперь же возьмем список игроков, попавших в первую и вторую защитные пятерки НБА. Данная награда вручается в результате голосования журналистов, игроков и тренеров ассоциации в конце сезона. В первую пятерку вошли: Руди Гобер (Юта, центровой), Энтони Дэвис (Новый Орлеан, форвард), Роберт Ковингтон (Филадельфия, форвард), Виктор Оладипо (Индиана, защитник), Дрю Холидей (Новый Орлеан, защитник). Вторая пятерка: Джоэл Эмбиид (Филадельфия, центровой), Дреймонд Грин (Голден Стейт, форвард), Эл Хорфорд (Бостон, форвард), Дежонте Мюррей (Сан-Антонио, защитник), Джимми Батлер (Миннесота, защитник). После чего отобразим на графике вышеперечисленных игроков (рисунок 2). Из графика можно увидеть, что большинство игроков являются лидерами и по статистике. Удивительно лишь попадание Дежонте Мюррея во вторую пятерку (рассчитанный защитный показатель 0.31 за 22 минуты игры, что чуть выше среднего).

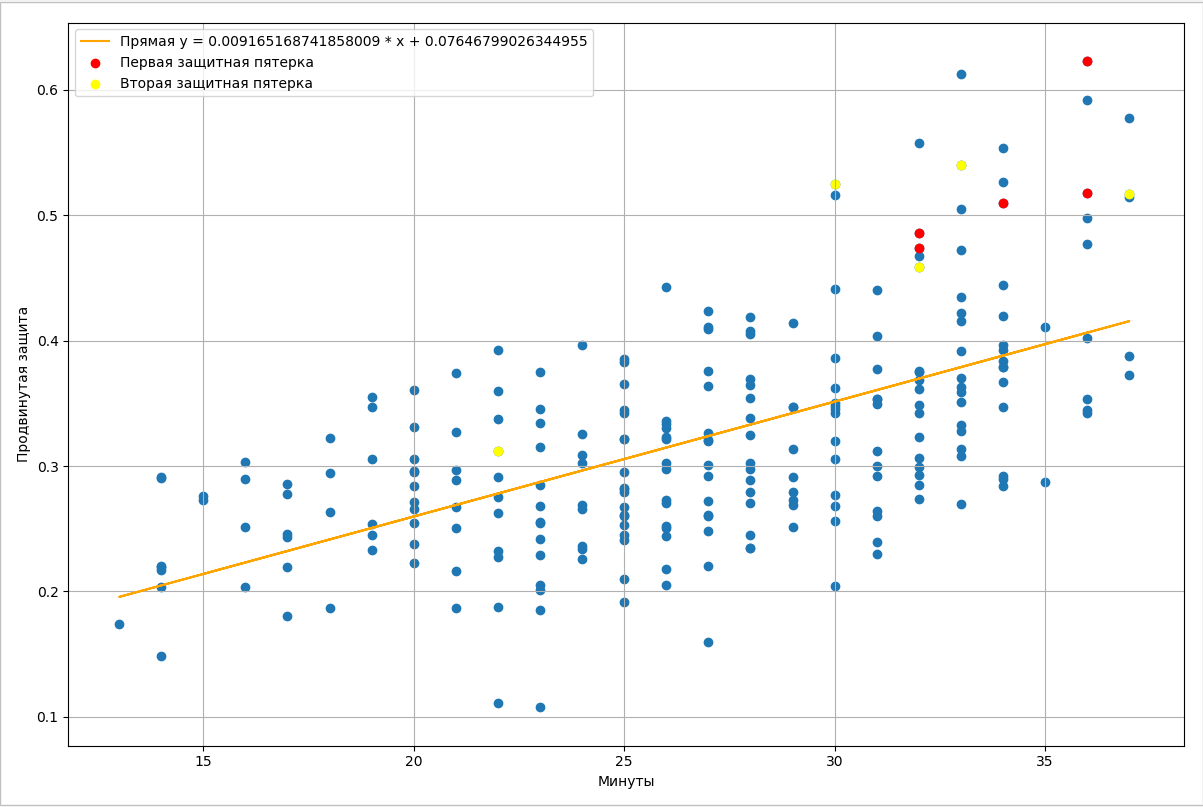


Рисунок 2 - Игроки 1 и 2 защитной пятерок NBA

В качестве альтернатив при выборе можно было рассмотреть Стивена Адамса (Оклахома, центровой - 0.6126 за 33 минуты в среднем), Бена Симмонса (Филадельфия, защитник - 0.5267 за 34 минуты в среднем) и Янниса Антетокунмпо (Милуоки, форвард - 0.5775 за 37 минут в среднем).

**Заключение**

Полученная модель позволяет оценить действия игроков Национальной баскетбольной ассоциации по совокупности 4 введенных в статистическую базу в сезоне 2017-18 показателей, а также зависит от снижения процентов попадания оппонентов. В результате получается новый и альтернативный подход к рассмотрению действий игроков при игре в обороне, затрагивающий максимальное количество статистических показателей.

**Список используемой литературы**

1. Сайт https://arxiv.org [Электронный ресурс]. – Режим доступа: <https://arxiv.org/pdf/1405.0231.pdf>
2. Сайт https://stats.nba.com [Электронный ресурс]. – Режим доступа: [https://stats.nba.com/stats/{endpoint}](https://stats.nba.com/stats/%7bendpoint%7d)
3. Сайт https://github.com [Электронный ресурс]. – Режим доступа: <https://github.com/seemethere/nba_py/wiki/stats.nba.com-Endpoint-Documentation>
4. Сайт https://academicjournal.ru [Электронный ресурс]. – Режим доступа: <https://academicjournal.ru/images/PDF/2018/Academy-7-34.pdf/Academy-7-34.pdf>